HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Expression and purification of inositol requiring enzyme 1 (IRE1)

Lamia Taha
Master's Programme in Genetics and molecular biosciences
Faculty of Biological and Environmental Sciences
University of Helsinki
2021

HELSINGIN YLIOPISTO

Tiedekunta – Fakultet – Faculty	Koulutusohjelma – Utbildningsprogram – Degree Programme
Faculty of Biological and Environmental Sciences	Master's programme in genetics and molecular biology

Tekijä – Författare – Author
Lamia Taha

Työn nimi – Arbetets titel – Title
Expression and purification of inositol requiring enzyme 1 (IRE1)

Oppiaine/Opintosuunta – Läroämne/Studieinriktning – Subject/Study track
Biochemistry and structural biology

Työn laji – Arbetets art – Level	Aika – Datum – Month and year	Sivumäärä – Sidoantal – Number of pages
Master's degree thesis	03/2021	50

Tiivistelmä – Referat – Abstract

The endoplasmic reticulum (ER) is an important organelle of the cell where a high number of proteins are synthesized and modified to obtain their final structure. Therefore, the ER stress, which is caused by accumulation of unfolded proteins in the ER, is not to be taken lightly since it could contribute to many diseases, such as cancer and neurodegenerative diseases. The response to the ER stress is the unfolded protein response (UPR), which is an adaptive system that helps in adjusting for increased folding needs within the ER. One of the main protein branches in the UPR is inositol requiring enzyme 1 (IRE1). IRE1 detects the status of protein folding inside the ER and initiates the UPR signaling pathway to achieve either normal folding status or cell death. The aim of this research was to express yeast IRE1 in *E.coli* and human IRE1 in insect cells, purify with affinity chromatography and study the IRE1's crystal structure with a small molecule modulator that could possibly enhance its activity. The protein was expressed successfully and purified with glutathione S-transferase (GST) tag, and the activity of the pure protein was determined. The structural studies were not fully completed since the absolute purity and yield that was necessary for crystallization was not achieved due to loss of protein during gel filtration and precipitation. Based on the results it is likely that the structure of the protein could be solved and further biochemical and structural studies with F10 are possible.

Avainsanat – Nyckelord – Keywords
ER, ER stress, IRE1, expression, purification, protein, GST

Ohjaaja tai ohjaajat – Handledare – Supervisor or supervisors
Ville Paavilainen, PhD, docent, and Shahid Rehan, PhD

Säilytyspaikka – Förvaringsställe – Where deposited
HELDA - Digital Repository of the University of Helsinki

Muita tietoja – Övriga uppgifter – Additional information

HELSINGIN YLIOPISTO

Tiedekunta – Fakultet – Faculty	Koulutusohjelma – Utbildningsprogram – Degree Programme
Bio- ja ympäristötieteen tiedekunta	Genettikan ja molekulaariset biotieteet

Tekijä – Författare – Author
Lamia Taha

Työn nimi – Arbetets titel – Title
Inositolia vaativan ensyymi 1 (IRE1) ekspressointi ja puhdistus

Oppiaine/Opintosuunta – Läroämne/Studieinriktning – Subject/Study track
Biokemia ja rakennebiologia

Työn laji – Arbetets art – Level	Aika – Datum – Month and year	Sivumäärä – Sidoantal – Number of pages
Pro gradu-tutkielma	03/2021	50

Tiivistelmä – Referat – Abstract

Endoplasminen retikkeli (ER) on tärkeä organelli solulle, missä suuri määrä proteiineja syntetisoidaan ja muokataan lopulliseen toiminnalliseen muotoonsa. Tämän takia ER:in stressiä, joka johtuu väärin laskostuneiden proteiinien kerääntymisestä ER:iin, ei pidä ottaa kevyesti, koska se voi vaikuttaa moniin sairauksiin, kuten syöpään ja eri neurodegeneratiivisiin sairauksiin. Täten laskostumattomien proteiinienvaste (UPR) on mukautuva järjestelmä, joka auttaa ER:ia sopeutumaan lisääntyneeseen laskostumistarpeeseen. Yksi vasteen pääproteiineista on inositolia vaativa entsyymi 1 (IRE1). IRE1 havaitsee proteiinien laskostumisen tilan ER:issa ja käynnistää UPR-signalointireitin, jotta saavutetaan joko normaali laskostumistila tai solukuolema.
Tämän tutkimuksen tarkoituksena oli tuottaa hiivan IRE1 proteiinia *E.coli*ssa ja ihmisen IRE1 hyönteissoluissa, puhdistaa proteiini affiniteettikromatografialla ja tutkia sen kiderakennetta pienen molekyylimodulaattorin kanssa, joka voisi mahdollisesti tehostaa sen toimintaa. Proteiini tuotettiin onnistuneesti ja puhdistettiin glutationi-S-transferaasi (GST)-merkinnän avulla ja eristetyn proteiinin aktiivisuus määritettiin. Rakenteellisia tutkimuksia ei suoritettu, koska kiteytymiseen tarvittavaa absoluuttista puhtautta ja saantoa ei saavutettu johtuen proteiinin menetyksestä geelisuodatuksen aikana ja saostumisen takia. Tulosten perusteella on todennäköistä, että proteiinin rakenne voitaisiin ratkaista ja biokemialliset ja rakenteelliset tutkimukset F10:llä ovat erittäin mahdollisia.

Avainsanat – Nyckelord – Keywords
ER, ER stressi, IRE1, ekspressio, puhdistus, proteiini, GST

Ohjaaja tai ohjaajat – Handledare – Supervisor or supervisors
Ville Paavilainen, FT, dosentti ja Shahid Rehan, FT

Säilytyspaikka – Förvaringsställe – Where deposited
HELDA - Helsingin yliopiston digitaalinen arkisto

Muita tietoja – Övriga uppgifter – Additional information

Table of contents

FIGURES

TABLES

Abbreviations

ER	Endoplasmic reticulum
UPR	Unfolded protein response
IRE1	Inositol requiring enzyme 1
PERK	PRK-like kinase
ATF6	activating transcription factor 6
F10	rapafucin
DNA	deoxyribonucleic acid
RNA	ribonucleic acid
mRNA	messenger RNA
rRNA	ribosomal RNA
tRNA	transfer RNA
elFs	eukaryotic initiation factors
Met	Methionine
Hsp70	Heat-shock protein 70
ATP	adenosine triphosphate
PDI	protein disulfide isomerase
PPI	peptidyl-propyl *cis-trans* isomerase
BiP	binding-immunoglobulin protein
RNase	ribonuclease
XBP1	X-box-binding protein 1
elF2α	eukaryotic translation initiation factor-2α
ATF4	activation transcription factor 4
PMDs	protein misfolding disorders
ERdj4	ER-localized J-protein 4
GST	glutathione S-transferase
His	histidine
Ni-NTA	Ni^{2+}- nitrilotriacetic acid
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
Sf21	Spodoptera frugiperda cell line 21
A_{280}	Absorbance at 280 nm
LB	Luria Broth
OD	optical density
IPTG	Isopropyl β- d-1-thiogalactopyranoside
rpm	revolutions per minute
MQ	Milli-Q water
AIM	Auto induction media
TBE	Tris-Borate-EDTA
DEPC	diethylpyrocarbonate treated water
PMSF	phenylmethylsulfonyl fluoride
PCR	polymerase chain reaction
PBS	Phosphate-buffered saline

1. Introduction

The endoplasmic reticulum (ER) is an important organelle in the cell, it has a continuous membrane with the nucleus and has many important functions. These functions are vital for the cell, such as synthesis, folding, transporting and modification of membrane and secretory proteins, in addition ER is also the main site of lipid biosynthesis. In the cell ribosomes translate proteins by combining amino acids, which will form a polypeptide chain. This chain will be inserted into the ER lumen (soluble secreted proteins) or inserted into the lipid bilayer (integral membrane proteins) through a membrane channel known as Sec61 in eukaryotes and SecY in eubacteria and archaea. If the chain is inserted or traversed into the membrane during translation to polypeptide chain the process is known as cotranslational translocation or if the polypeptide chain is completed before insertion into the Sec61 channel it is known as posttranslational translocation (Crowley et al., 1993; Mackinnon et al., 2014; Osborne et al., 2005). Inside the ER, proteins are then folded and modified with the help of specific chaperone proteins, after which the mature proteins are then transported to their final destination. However, due to some changes in the cell's environment or mistakes in the genetic information will cause the proteins to be misfolded and start to accumulate inside the ER lumen. This causes the ER to undergo to a state of discomfort that is known as "ER stress".

The survival of the cell depends on the correct functions of its organelles, homeostasis describes the process that is required to maintain stable conditions necessary for survival. In ER the survival is dependent of proteostasis, the stable conditions of protein folding and function. There is an signaling pathway that detects and monitors the ER status, it is called the unfolded protein response (UPR). This response will try to restore the homeostasis in the ER, by activating its main UPR proteins; IRE1, PERK and ATF6 (in yeast only IRE1), which will either increase the folding capacity of the ER or activate the expression of UPR components. Nevertheless, sometimes the folding capacity cannot be restored, which will then lead to cell death by apoptosis.

IRE1, as one of the main proteins in the UPR, has been found to have connections to many health issues, such as cancer, Parkinson's disease, and Alzheimer's disease (Wang and Kaufman, 2016). Therefore, modulating the IRE1 function to benefit in therapeutical solution to said diseases is something to be further investigated with gene therapy, but small molecule approaches are needed. The structure of IRE1 is mostly known (Korennykh et al.,

2008; Korennykh et al., 2011), but the structure with a small molecular modulator rapafucin F10 is yet to be discovered. This undiscovered structural form could be the key that would lead to new approaches in finding solutions for treating various diseases.
In this project the main goal is to express, purify and use x-ray crystallography to understand how the small molecule modulator, rapafucin F10, binds to IRE1 and how the binding affects it activity and structure.

1.1 Protein synthesis and folding

Protein synthesis begins with deoxyribonucleic acid (DNA) that is transcribed to ribonucleic acid (RNA) in the nucleus after which the RNA is transferred into to the cytosol of the cell. There are different kinds of RNAs, such as messenger RNA (mRNA), ribosomal RNA (rRNA) and transfer RNA (tRNA). The mRNAs nucleotide sequence, 3 nucleotides make a codon that corresponds to one amino acid out of 20 (22 in prokaryotes), is used as a template by ribosomes and by tRNA, which needs to be aminoacylated to be in its active form, to synthesize polypeptide chains (Litwack, 2018). The aminoacylated tRNA has a structure that resembles a cloverleaf, and in one loop there is an anticodon for the mRNAs different sequence possibilities, and on the other end it has an amino acid bound to it by a aminoacyl-tRNA synthetase (every amino acid has its own aminoacyl-tRNA synthetase) that corresponds the anticodon. The ribosome, which consists of two subunits 40S and 60S, detects and binds to the mRNA with the help of many different eukaryotic initiation factors (eIFs). First the smaller 40S subunit binds to mRNA along with some eIFs and Met-tRNA, which is the first amino acid in mRNA also known as initiation or starting codon (Kozak, 2002; Preiss and W. Hentze, 2003). Then the 60S subunit of the ribosome bind to the 40S, forming the 80S ribosome (Preiss and W. Hentze, 2003), which will then synthesize the polypeptide chain by reading the mRNA codons and combining it with the corresponding anticodon-tRNAs. Thus, the polypeptide chain will grow starting from its N-terminus until the stop-codon, which will terminate the synthesis of the polypeptide by having a release factor binding to it. The release factor frees the complete polypeptide chain (C-terminus end) and the ribosome subunits are released from the mRNA (Litwack, 2018).
Once the polypeptide chain is freed from the ribosome, it starts folding into its correct form mostly with the help of chaperones, which is determined by the amino acids. There are 20 different amino acids, they share the same carboxylic acid group and amino group, but the side chain that is bound to the C_α in the middle is where they differ. Some are for example

hydrophobic, hydrophilic, neutral, contain rings or sulfur (only cysteine). Protein folding is driven by the interactions of the side chains, the polypeptide main chain and its backbone, where it will hide away from the surface of the protein the hydrophobic residues and form a few different kinds of salt bridges and disulfide bonds with the help of polypeptides (Hebert and Molinari, 2007).

The polypeptide chain is known as primary structure, but there is also secondary structure which is when the polypeptide chain forms α-helixes and β-sheets. There are sometimes more than one helix or sheet in a proteins secondary structure these then forms the tertiary structure together, also known as domains. Some large proteins need multiple domains to function and this collection of multiple domains is known as tertiary structure.

Polypeptides have also help from other proteins to achieve and maintain their correct structure, these proteins are known as chaperones and can be found where proteins are being synthesized and going through posttranslational modifications in the cell (Ellis, 1987; Hartl, 1996; Hebert and Molinari, 2007). Chaperones bind to the polypeptide chain and assist it to gain its correct structure and therefore lower the possibility of incorrect structures (Ellis, 1987). The heat-shock protein 70 (Hsp70) and chaperonin family are the most known chaperone proteins. Chaperonin helps the folding happen in a cavity of the chaperonin itself and the Hsp70 uses a method that requires adenosine triphosphate (ATP) hydrolysis (Hartl, 1996).

Chaperones are not the only thing that are helping proteins to gain their functional structure, there are also foldases, protein disulfide isomerases (PDI) and peptidyl-propyl *cis-trans* isomerases (PPI). Foldases are enzymes that increases the speed of the folding reaction, PDI is an enzyme that helps cysteine residues to form a covalent bonds to other cysteine residues and PPI is a class of enzymes that catalyzes the isomerization of peptidyl-propyl bonds (Hebert and Molinari, 2007).

Even though there are many proteins that help polypeptides into their correct structure, the protein could still fail to acquire their correct fold due to mutation or environmental and cellular stresses (Wang and Kaufman, 2016).

1.1.1 Protein translocation into the ER

The ribosomes in eukaryotic cells are either free in cytosol and in mitochondria or bound to the ER membrane. The nuclear-encoded mitochondrial, peroxisomal and cytosolic proteins

are synthesized in cytosol by cytosolic ribosomes. Most of membrane and secretory proteins are being synthesized by ribosomes that are bound to the ER membrane (Hebert and Molinari, 2007; Osborne et al., 2005).

The ribosomes combine the amino acids into polypeptide chains, which is being inserted into the ER through a transmembrane complex known as Sec61 translocon on the ER membrane (Gemmer and Förster, 2020; Johnson and van Waes, 1999).

The nascent polypeptide chains go through co- and post-translational modifications and start folding with the help of various chaperones. The Hsp70-type chaperone in the ER lumen, also known as BiP (binding-immunoglobulin protein), is the most abundant chaperone protein inside the ER (Adams et al., 2019). Chaperone activity and modifications made to proteins, such as N-linked glycosylation, formation of disulfide bond, are important for correct protein folding. The addition of oligosaccharides is one way to track the protein folding in the ER (Wang and Kaufman, 2016).

When the proteins fail to obtain their final structure, ER-associated degradation (ERAD) is initiated (Adams et al., 2019; Hampton, 2002). This is a process, that sends the misfolded protein from ER lumen to the cytosol of the cell. In the cytosol the misfolded protein will undergo ubiquitylation by ubiquitin ligases and send for further degradation by the proteasome (Meusser et al., 2005).

Different types of cells, such as plasma cells, or insulin-producing cells, where the ER translates different sets of proteins that constitute to the secretome. Thus, the intake of nascent polypeptide chains differs widely from cell to cell, this could disturb the balance of the folding capacity of the ER, therefore the amount of misfolded proteins will grow. The growth of unfolded proteins is toxic to the cell, which causes the phenomena known as “ER stress” (Adams et al., 2019). The homeostasis in ER can be restored through a signaling network that observes the protein folding status in ER and responds to it, known as the unfolded protein response (UPR) (Wang and Kaufman, 2016) that also generates ER size alteration (Schuck et al., 2009; Walter and Ron, 2011).

1.2 ER stress and the unfolded protein response

The UPR updates the nucleus of the cell about the protein folding condition in the ER lumen and the gene transcription attempts to adjust the ER protein folding efficacy depending on the UPR signal that riggers transcriptional response (Hetz, 2012). In the mammalian UPR signaling network there are three main proteins, that are activated; PRK-like ER kinase

(PERK) (Harding et al., 1999), inositol requiring enzyme 1 (IRE1) (Cox et al., 1993), and activating transcription factor 6 (ATF6) (Haze et al., 1999). All have an ER-lumen domain, transmembrane domain, and a cytosolic domain.

The activation of UPR (Figure 1) is induced by oligomerization of IRE1 protein, also upon oligomerization IRE1's kinase and ribonuclease (RNase) domains are activated (Prischi et al., 2014) When active the RNase domain splices the X-box-binding protein 1 (XBP1) mRNA, the spliced XPB1 is a transcriptional activator (HAC1 in yeasts) (Calfon et al., 2002; Cox and Walter, 1996). Which leads to an increase in the expression of the potent transcriptional activators and then it will upregulate the transcription of needed chaperones, lipids and ERAD proteins (Ron and Walter, 2007). Furthermore, there is a process known as regulated IRE1 dependent decay (RIDD), that results in mRNA decay that are targeted for ER membrane, which will further aid in decreasing the protein demand in the cell (Hollien and Weissman, 2006).

PERK has a similar structure as IRE1, which leads to similar activity. The kinase domain in PERK phosphorylates the eukaryotic translation initiation factor-2α (eIF2α), this leads to ribosome inhibition and depletion of proteins translation in the whole cell (Harding et al., 1999). However, even though this leads to reduced protein translation in the ER, the genes of activation transcription factor 4 (ATF4) are being upregulated, to respond to the cellular stress damage with antioxidants (Vattem and Wek, 2004).

The third UPR protein ATF6 differs in its function, once activated it will leave the ER membrane and go to the Golgi apparatus, where site specific proteases (S1P and S2P) will cleave the luminal domain from the cytosolic domain of the protein itself. The cytosolic domain transits to the nucleus and begins upregulating UPR destined and ER chaperone genes (Haze et al., 1999; Shen et al., 2002).

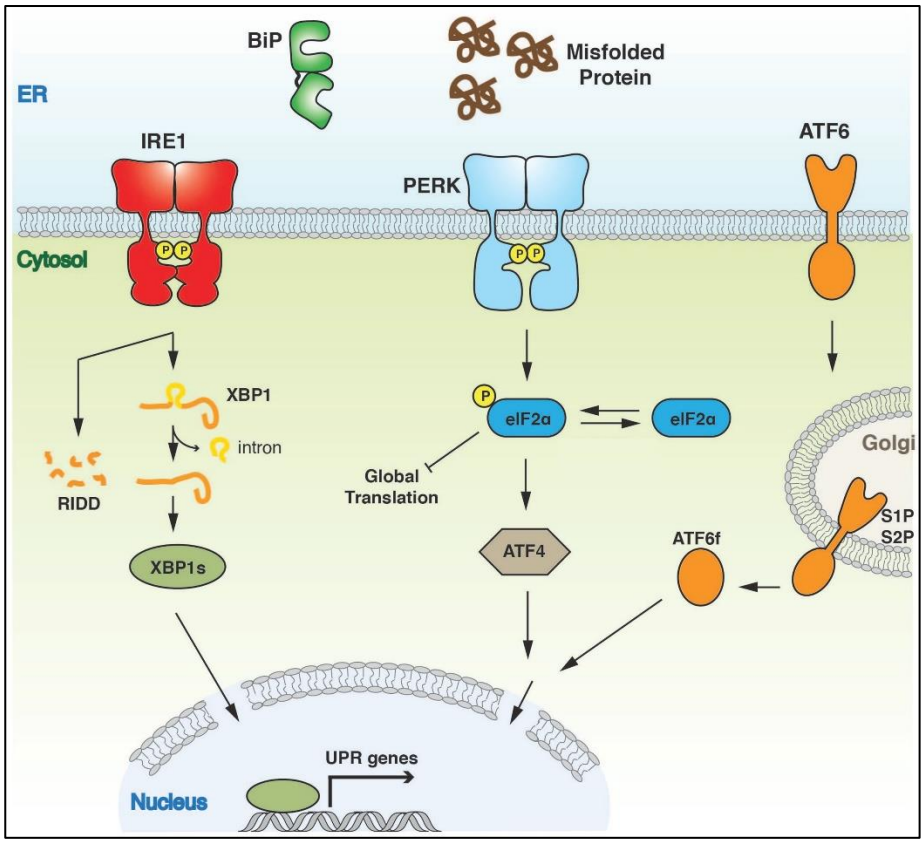

Figure 1 – UPR signaling pathways. In the UPR signaling pathway IRE1 splices XBP1 to increase the expression of transcription activators to upregulate UPR genes in nucleus. RIDD results in mRNA decay, that will decrease protein demand in cells. PERK phosphorylates elF2 α, which leads to ribosome inhibition and depletion of protein transition in the ER. Also, ATF4 are being upregulated to respond to cellular stress. Once activated ATF6 goes to Golgi apparatus and to be cleaved by the S1P and S2P, the cytosolic domain will transit to nucleus to upregulate URP genes. Only the IRE1 pathway is conserved from yeast to humans. Image from Adams et al., 2019.

1.2.1 Disease relevance of UPR

The demand of protein synthesis and secretion is different in different type of cells, for example the acinar and islet β-cells are specialized for polypeptide secretion in the pancreas. Therefore, there are cells that require higher efficiency of proteins synthesis and folding than other cells. Nevertheless, in the ER the protein folding success is highly dependent of its environment in and outside of the ER, which means that almost any kind of a change in the environments will make a difference. These changes could be for example limited energy availability, increasing demand of protein synthesis, deficient autophagy, hypoxia and inflammatorial changes. The ER environment is different to the cytosolic, which is oxidic in comparison, to the cell. This and the specific Ca^{2+} concentration in ER are essential for folding and chaperone activity (Wang and Kaufman, 2016). When the protein

folding does not occur accordingly, then the misfolded proteins will accumulate in the ER. This has been a characteristic sign in many neurodegenerative diseases as one of the faults that could be the cause of the condition itself (Cabral-Miranda and Hetz, 2018). Such diseases as Alzheimer's disease, amyotrophic lateral sclerosis, prion-related disorder, and Parkinson's disease, are being called protein misfolding disorders (PMDs) (Cabral-Miranda and Hetz, 2018; Soto, 2003).

In addition, to protein folding the ER has a role in the biogenesis of peroxisomes and autophagosomes. The UPR is responsible to eliminate the cells with apoptosis that were not able to regain the protein folding homeostasis in ER (Hetz, 2012).

1.3 IRE1 protein

1.3.1 IRE1 signaling and function

The IRE1 luminal domain is the part of the IRE1 that identifies the ER stress, it will then lead to activation of the protein. There are some suggestions of how the IRE1 identifies the ER condition, the direct association, competition, and allosteric model. Here we will go through these models briefly (Adams et al., 2019).

In the direct model it is believed that the luminal domain binds to the misfolded proteins directly, which then will activate the URP signaling pathways. There is a proposal that the luminal domain has a unique fold and is sufficient to control the activation of IRE1 due to its complex structure. The crystal structure of luminal domains showed a binding structure that is similar to the major histocompatibility complex (MHC) fold structure (Credle et al., 2005). In this model the BiP does not have a main role but more a role in adjusting the activity of IRE1 (Kimata et al., 2004).

In the competition model BiP prevents dimerization of IRE1 and thus represses the UPR signaling pathway. BiP binds to IRE1 in a specific chaperone-substrate interaction type, which is moderated by ER-localized J-protein 4 (ERdj4). This interaction forms a repressive complex that leads to BiP ATPase activity causing ERdj4 to dissociation. This dissociation makes IRE1 form monomers which stops the UPR signaling. But in ER stress misfolded proteins occupy BiP and ERdj4, which frees IRE1 to be activated (Amin-Wetzel et al., 2017).

In allosteric model BiP' s ATPase domain is bound to IRE1 and is released from it when misfolded proteins bind to BiP' s substrate-binding domain. The release of causes a

conformational change in IRE1 that activates the UPR signaling, making BiP the sensor of the accumulation of the misfolded proteins (Carrara et al., 2015; Kopp et al., 2018).
Once activated the IRE1 oligomerizes and the cytoplasmic kinase domains are near to other kinase domains in an orientation known as face-to-face. Due to this close locations to each other the kinase domains are autophosphorylated, which leads to activation of RNase domain to splice XBP1 mRNA (HAC1 in yeast) (Ali et al., 2011; Prischi et al., 2014; Ron and Walter, 2007).

1.3.2 IRE1 structure

The IRE1 human and yeast structures have been obtained throughout the years (Ali et al., 2011; Credle et al., 2005; Joshi et al., 2015; Korennykh et al., 2008; Korennykh et al., 2011; Lee et al., 2008; Tirasophon et al., 1998; Zhou et al., 2006). They all state that IRE1 has three domains: luminal, transmembrane, and cytosolic domain. The N-terminal end of IRE1 is in the ER lumen, making it the luminal domain, followed by the transmembrane domain which passes the ER membrane once. In the cytosolic domain with the C-terminal are the kinase and RNase domain (Figure 2).
The cytosolic domain has two structural forms, in one form the kinase domain is in the face-to-face formation that is needed for the autophosphorylation, which then further leads to its other form known as back-to-back formation that is assumed to be the RNase active form (Adams et al., 2019). Also, when the kinase domain binds to ATP-binding inhibitors it can have effects on the RNase activity (Concha et al., 2015; Papa et al., 2003).

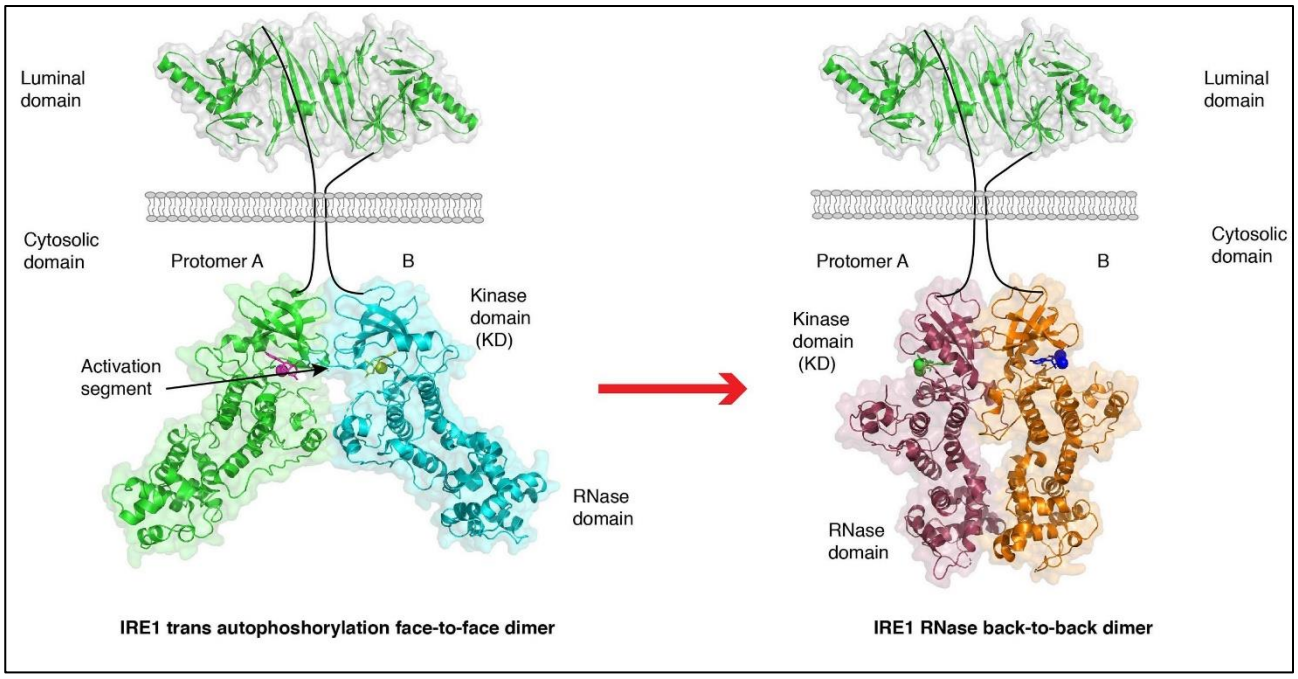

Figure 2 –Overall structure of IRE1. The IRE1 cytosolic domain in dimeric mode, when the configuration is face-to-face the kinase domain is active. When the configuration is back-to-back then the RNase domain is active. Kinase domains binding to ATP-binding inhibitors can activate the RNase activity. We hypothesize that F10 would activate IRE1 RNase by binging to the kinase domain. Image from Adams et al., 2019

1.4 Rapafucin F10

The IRE1 protein in the UPR signaling pathway is important in many aspects for cells survival and wellbeing, therefore there have been research on different molecules to see if they could enhance or inhibit IRE1 function. Different regulations of IRE1 signaling have been reviewed by (Hetz et al., 2020)

Rapafucin F10 is a rapamycin-like small molecule that has recently been identified in a high content screen as an activator of XBP1 splicing. The rapafucin is hybrid macrocycle synthesized from rapamycins and FK506, which are natural macrocyclic compounds. These compounds can bind to specific proteins and form complexes with their target proteins (Guo et al., 2019).

2. Aims

The aim of this project was to express both yeast and human IRE1 protein for biochemical studies and structural biology investigations. Yeast IRE1 protein was to be expressed in *Escherichia coli (E.coli)* and human IRE1 in baculovirus infected-insect cells. Then the protein was to be isolated and purified from the cells with affinity chromatography and used for structural work.

Yeast IRE1 is GST-tagged (Korennykh et al., 2008) that was used for isolation and purification, with glutathione beads that bind to GST-tagged proteins. The human IRE1 is 6His-tagged (Harnoss et al., 2019) and was to be purified with Ni-NTA beads that bind to 6His-tagged proteins.

The ultimate purity was to be achieved by using affinity chromatography and gel filtration and when the yield was high enough crystallization of both human and yeast constructs was to be attempted. The crystallization was to be attempted in both ways, sitting and hanging drop. The F10 will be added to some before crystallization and to some after the protein is crystallized. This was to allow us to optimize our chances to obtain a co-crystal structure of either yeast or human IRE1 bound to F10. The aim was to get the chance to solve the X-ray structures of human and yeast IRE1 with and without F10.

Another aim was to find out if the expressed, purified human and yeast IRE1 are active, by performing an mRNA cleavage assay with HAC1 and XBP1. Additionally, to see if F10 has any effects to the activity at various concentrations. Hence it could serve as an IRE1 protein activator and therefore as a model for pharmaceutical solutions.

3. Materials and Methods

3.1 Materials

The culture media, plates and most of the solutions were purchased from the Institute of Biotechnology Media Kitchen. The Amicon Ultra concentration tubes were purchased from Merk Millipore Ltd. The SDS-PAGE gel (Mini-Protean TGX Stain-Free Precast Gel, Bio-Rad laboratories, Inc.) and the RNA gel (Mini-Protean TBE-Urea Precast Gel, Bio-Rad laboratories, Inc.) were run with PowerPac Basic (Bio-Rad laboratories, Inc.). The gels were imaged with Bio-Rad Gel Doc™ EZ Imager. Standard Agarose – Type LE (Bio Nordika Oy) was used to prepare 1% agarose gels. The insect cells Sf21 were purchased from Invitrogen. For western blotting the blocking buffer (EveryBlot Blocking buffer, Bio-Rad laboratories, Inc.), pads (Trans-Blot® Turbo™ Mini-size Transfer Stacks, Bio-Rad laboratories, Inc), membrane (Trans-Blot® Turbo™ Mini-size nitrocellulose, Bio-Rad laboratories, Inc.) and transferring from gel to blot Trans-Blot® Turbo™ Transfer system was used.

3.2 Yeast IRE1-GST

The yeast IRE1-GST constructs PWD 1475 and PWD 1477 were a kind gift from Professor Peter Walter (UCSF). These two constructs vary in the length of their amino terminal end, one is residues 673-1115 and other 641-1115 (see APPENDIX 1). But it is not yet clear which is which. Both are in pGEX-6P-2 vector, which contains the GST on the amino terminal end with a PreScission protease cleavage site (Korennykh et al., 2008).

3.2.1 Small scale expression test of constructs 1475 and 1477

The plasmids of both yeast IRE1 constructs were purified using the plasmid purification kit (MACHEREY-NAGEL GmbH & Co). Absorbance at 260nm (A_{260}) of the plasmids was measured (1475 = 1335.6 ng/µl and 1477 = 1260.1 ng/µl) with nanodrop (NanoDrop 2000c, Spectrophotometer, Thermo Scientific). Both plasmids were transformed to BL21 RIPL *E. coli* cells (Agilent Technologies, Inc.) using heat shock method. Briefly, cells were first thawed on ice and 15 µl of purified plasmids were added. The cells were incubated on ice for 25 min, then gave 1-minute heat shock at +42 °C after which they were incubated on ice

for 1 min. Then 500 µl of Luria Broth (LB)-media was added and incubated for 1h while shaking at +37°C. The cells were plated on LB-plate that contained ampicillin, chloramphenicol, streptomycin and spectinomycin (last three antibiotics required for RIPL cells), and were grown overnight in +37°C. From these plates 5 colonies were transferred to 10 ml LB-media liquid culture with the same antibiotics to grow overnight shaking in +37°C. The next day the overnight growth was diluted by taking 100 µl of the overnight culture to fresh 10 ml LB-Media with antibiotics, grew them until the OD_{600} (optical density, measured with Eppendorf BioPhotometer using disposable cuvettes) was 0.5, then induced it with 0.4 mM IPTG. The OD_{600} was measured only from one tube (1477). The induced liquid culture was grown for 2h in +20°C. 1 ml samples were taken from both before and after inducing for running the SDS-PAGE gel. The preparation of the samples for gel was first done by centrifuging them in benchtop centrifuge (Heraeus FRESCO 21 centrifuge; Thermo Scientific; 14 000 rpm, 5 min) and removing the supernatant. The pellets were resuspended to 1xPBS with 80 U DNase (New England biolabs Inc.) and left incubating 10 min in room temperature. Then samples were freeze-thawed 4 times (liquid nitrogen and waterbath at temp +32°C) and sonicated for 30 min in waterbath sonicator (Ultrasonic cleaner, BRANSON a SmithKline company, B-52, 248 Watts, 50/60 Hz), after which added 5x laemmli buffer to the sample and run in SDS-PAGE gel.

3.2.2 Small scale expression of yeast IRE1 construct 1477

The expression was done in the same way as before in section 3.2.1. The induction was done with a higher concentration of IPTG (1 mM) and the induced liquid culture was grown for 4 hour in +20°C. 1 ml samples were taken from both before and after inducing and normalized to OD_{600} = 1 for running the SDS-PAGE gel.

3.2.3 Large scale expression of yIRE1-1477

From previous LB-plate 5 colonies were grown overnight shaking (+37°C) in 150 ml LB media, with ampicillin, chloramphenicol, streptomycin and spectinomycin. 6l LB media was inoculated with overnight culture containing all the antibiotics. Cells were grown shaking +37°C until the OD_{600} reached 0.8 and then induced with 0.8 mM IPTG (5 ml). After 4h of inducing at room temperature (+22°C), the culture was centrifuged at 6000 rpm 30 min 4°C (Sorvall™ RC 6plus centrifuge, Thermo Scientific) on rotor F9-4x1000y (FIBERLite, Piramoon Technologies Inc.). Supernatant was discarded and pellets were flash frozen in

liquid nitrogen and stored at -80°C until further analysis. Before and after inducing with IPTG, 1 ml samples were taken and prepared the same way as described in section 3.2.1 for running SDS-PAGE gel.

3.2.4 Expression of yIRE1-1475 with carbenicillin

The expression of 10 ml of the construct 1475 was prepared as was in section 3.2.1 with antibiotics, but instead of ampicillin carbenicillin was used because it is more resistant to degradation. Cells were grown overnight shaking +37°C. Measured the next day OD_{600} which was 2.5 and diluted it to 0.2 and let it grow until OD_{600} reached 0.5. Added 0.8 mM IPTG and let it grow 4h in +22°C. Before and after inducing with IPTG 1 ml samples were taken and prepared the same way as described in section 3.2.1 for running SDS-PAGE gel.

3.2.5 Scaling up expression of yIRE1-1477

Prepared 25 ml of LB-media with antibiotics as was in section 3.2.1 and added 5 colonies of BL21 RIPL cells and let it grow overnight shaking +37°C. Prepared the next day 1l LB-media was inoculated with overnight culture containing all the antibiotics and let it grow in +37°C until the OD_{600} reached 0.5. Then added 0.8 mM IPTG and let it grow for 4h in +22°C. Collected the cells by centrifuging at 6000 rpm 30 min +4°C (Sorvall™ RC 6plus Centrifuge, Thermo Scientific) on rotor F9-4x1000y (FIBERLite, Piramoon Technologies Inc.) and froze the pellet in -80°C. Before and after inducing with IPTG 1 ml samples were taken and prepared the same way as described in section 3.2.1 for running SDS-PAGE gel.

3.2.6 Test purification yIRE1-1477

For test purification, pellet from one liter of cells was thawed and resuspended in buffer A (20 mM HEPES (pH 7.5), 150 mM KCl, 10 mM $MgCl_2$, 0.5 mM DTT, 10% (v/v) glycerol) with 1% Triton-X (MP Biomedicals LLC), 1 mM PMSF and protease inhibitor tablet (Pierce™ Protease Inhibitor Tablets, EDTA-Free, Thermo Scientific). Then lysed the cells with emulsiflex (Emulsiflex-C3, AVESTIN, Inc) for 20 min, 10 000 psi, and then the lysate was centrifuged 18000 rpm 30 min +4°C (Sorvall™ RC 6plus centrifuge, Thermo Scientific) on rotor F21-8x50y (FIBERLite, Piramoon Technologies Inc.). Meanwhile prepared Glutathione Sepharose 4B beads (GE Healthcare Bio-Sciences AB) by washing them with MQ-water and then equilibrated with the same buffer as was used for resuspension of pellets. After

centrifugation the clear lysate was collected and added to them the equilibrated beads, this mixture was left rotating for 1h in cold room. Then the mixture was centrifuged at 1000 xg 5 min (Eppendorf Centrifuge 5810 R, Eppendorf AG) in rotor A-4-62 and the flowthrough was collected through a column. Then 10 ml of buffer A was added to the beads in column and incubated for 10-15 min on ice, then the wash 1 was collected. This was repeated 3 more times to collect washes 2, 3 and 4. Next, 10 ml of buffer B (20 mM HEPES (pH 7.5), 0.5 mM DTT) was added to the beads and incubated just as had been done previously to collect the low salt wash and then 10 ml of buffer C (20 mM HEPES (pH 7.5), 0.5 mM DTT, 500 mM KCl) was added to collect high salt wash. Then added 10 ml of buffer A with 10 mM ATP to the beads, to get rid of chaperones, and collected the ATP wash. Then added 10 ml of buffer D (50 mM HEPES (pH 7.5), 150 mM KCl, 1 mM EDTA, 0.5 mM DTT) with 60 µl of protease C3 (Novagen, EMD Millipore Corp) and left rotating in cold room overnight in a falcon tube. Collected the elution the next day and run SDS-PAGE gel with 10 µl samples of each step. Based on the results of the SDS-PAGE gel, washed the beads with buffer D, then prepared 5 ml of buffer D with 15 mM glutathione (L-Glutathione reduced, Sigma-Aldrich Oy) and added it to the beads 1 ml at a time and collected 4 elutions. The absorbance at 280 nm of the elutions and of the overnight elution sample was measured. According to the absorbance results the overnight samples was concentrated and measured the A_{280} of them and run an SDS-PAGE gel with all these elutions and flowthrough of the concentrated sample. Pooled all elution (except the overnight elution that was concentrated) after collecting the 5^{th} elution, added 60 µl of protease C3 and dialyzed them in 1l of buffer D overnight. The sample was then collected and added to the beads that were washed with buffer D. Incubated for 1h, while mixing every 15 min with pipet, and collected the sample through the column. Measured A_{280} of the sample, which should be yIRE1 without GST. Then added glutathione containing buffer D to the beads to collect yIRE1-GST sample. Both protein versions were then concentrated using the concentration tubes with 30K cutoff filter (Centrifugal Filters, Amicon Ultra – 15) with centrifuging the samples in the tube at 4000 rpm 2-3 min in centrifuge (Eppendorf Centrifuge 5810 R, Eppendorf AG). Then 20% glycerol was added to the samples and divided into smaller portions and placed in -80°C. Calculated the concentrations with A_{280} values, the extinction coefficient for yIRE1-1477= 40340 and yIRE1-1477-GST= 69680, the concentration of yIRE1-1477 = 3.03 µM and of yIRE1-1477-GST = 8.64 µM.

3.2.7 Autoinducion media (AIM) testing with 1475 and 1477

Due to the slow growth of the construct 1475 autoinduction media (Grabski et al., 2005) was tested to see if the expression of 1475 could reach the expression level of 1477. Both constructs, yeast IRE1-GST 1475 and 1477, were grown in 600 ml in autoinduction media (AIM−terrific broth base including trace elements, Formedium^{TM}) with same antibiotics as previously was done in section 3.2.1 from freshly transferred colonies from previous plate. They were grown in +30°C while shaking. The liquid culture was centrifuged, and 1 ml pellet was treated as was done before in section 3.3.2 and run on the SDS-PAGE gel.

3.2.8 Purification of AIM expressed 1475 and 1477

The purification was done mostly as was done with 1l test purification in section 3.2.6 with the exception that the binding of the IRE1-GST was done overnight rotating in the cold room. And since the beads are re-usable used the same beads after washing them thoroughly and divided them between the two constructs. From each washing step 10 µl samples were taken for running the gel. After performing all the washes, the IRE1-GST construct was eluted from the beads with 15 mM glutathione in buffer D. In total 3 elution was collected and later combined the samples that according to the SDS-PAGE gel contained IRE1-GST. The combined sample was concentrated with 30K cutoff filter by centrifuging and during concentrating exchanged the previous buffer to buffer A, by using 30 ml of it. At the end the volume was around 1.1 ml (1475) and 0.7 ml (1477). Measured the absorbance and calculated the concentration of the proteins, the 1477 was 8.78 µM and 1475 was 0.98 µM. Both were then aliquoted to 100 µl /tube with 20% glycerol and stored flash frozen in liquid nitrogen in -80°C.

3.2.9 mRNA cleavage reaction

The HAC1 mRNA was provided by PhD student Vanessa Bensch from the same lab and tested for the activity of the yIRE1 constructs. The 20 µl reaction contained 400 nM of protein, 3ul mRNA, 1 mM ATP and rest buffer E (20 mM HEPES (pH 7.5), 10 mM Magnesium acetate, 50 mM potassium acetate, 1 mM DTT). Left incubating in +30°C for 3 hours, and then added the RNA loading dye. The samples were then kept for 5 min at +65°C, after which they were immediately placed on ice and then run the RNA gel 180V with

0.5xTBE buffer with DEPC-treated water. The gel was stained with SYBR Gold (SYBR Gold Nucleic Acid Gel Stain) for 10 min and imaged.

3.2.10 Large scale purification of yIRE1

The pellet of 5l was first thawed and resuspended with 35 ml buffer A containing 1% Triton-X, 1mM PMSF and 1 tablet of protease inhibitor. After resuspension the cells were lysed with emulsiflex for 10 min and then centrifuged for 30 min at speed 18 000 rpm. The cleared lysate was collected and added to 10 ml glutathione sepharose beads for overnight binding. Collected the flowthrough by centrifuging 5 min 1000 xg and started washing the beads as was done in section 3.2.3, but instead of using 10 ml of buffers in washes 25 ml was used. Also, all steps were collected by centrifuging at 1000 xg 5 min. After the washes eluted the protein from the beads by placing them in a 20 ml column with buffer D containing 15mM glutathione, collected 5 ml elutions by incubating 10 min between each elution until had collected 8 elutions in total. Measured their A_{280} with nanodrop and combined them trusting the nanodrop concentration and concentrated them from 40ml to 5ml. The A_{280} of the concentrated sample was 9.123 mg/ml. From every wash step 10 µl samples were taken and then run on gel. Added protease in ratio 1:10 (protease to protein) to 5 ml protein sample to cleave off the GST and left it rotating overnight. Run on gel to see if the cleavage has occurred and added the sample to glutathione beads for 90 min while rotating to capture the free GST and uncleaved IRE1-GST. Then collected through the column 17 ml in total of elution which had the A_{280} of 2.287 mg/ml. To this sample added Ni-NTA (Pierce™ HisPur™ Ni-NTA Superflow Agarose, Thermo scientific) beads to capture the protease for 30 min. Collected the protease free sample by centrifuging it at 1000 xg 5 min and concentrated the rest of the sample to 3 ml (A_{280} = 6.655 mg/ml).

3.2.11 Gel filtration

The 3 ml sample was injected into size exclusion (HiLoadTM 16/60, Superdex 75 prep grade, Pharmacia XK16) gel filtration (ÄKTA purifier, amersham pharmacia biotech) in buffer 20 mM HEPES (pH 7) (20°C), 500 mM NaCl, 2 mM DTT, 5% glycerol and the collection of 1 ml fractions started at 40 min. During the gel filtration 4 peaks were observed and then collected to see which of them is protein by running an SDS PAGE gel with the corresponding fractions. Once the corresponding fraction was found, concentrated it from 11 ml to 500 µl. Meanwhile, tried to see it there still was some of the protein left in the

glutathione sepharose 4B column by collecting 6 more elutions of 5 ml with glutathione containing buffer and run the samples on a gel with the concentrated peak fraction from the gel filtered column.

3.2.12 12l expression

Transformation of new BL21 RIPL *E. coli* cells was redone as in section 3.2.1 only with yeast IRE1-GST-1477 construct to make a new expression of the yeast IRE1-GST. From the new plate 5 colonies was picked for the 500 ml overnight culture with accurate antibiotics. For 12l expression 500 ml of 2x LB-media was used and diluted with 500 ml MQ-water, to which added 1 ml ampicillin, 1 ml chloramphenicol, 3 ml streptomycin-spectinomycin mixture and 40 ml of overnight culture. Grew the liquid cultures shaking in +37°C until the OD_{600} was 0.5, induced the growth with 5 ml IPTG and lowered the temperature to +22°C. Samples has been collected to check the expression on gel before and after inducing on gel like was done in section 3.2.1. Harvested the 12l liquid culture by centrifuging 4 bottles at once (1l/bottle) at 6000 rpm 30 min, at the end there were 4 bottles that had pellets of 3l culture in them, kept them frozen -80°C until purification.

3.2.13 12l purification

Thawed the pellets and resuspended them in 45 ml of buffer A with 1% Triton-X, 1 mM PMSF and a tablet of protease inhibitor. Lysed the cells for 10 min on the emlusiflex and centrifuged 18000 rpm 30 min. Used the same beads as before and new ones (all of it), in total had around 20 ml beads, kept them in separate falcon tubes for cleaning and equilibrating in buffer A. Added the clear lysate to the beads (divided to two tubes) and let them rotate for 2h. Collected a small sample from the flowthrough and run a gel to see if the binding has been sufficient. Then continued the binding by rotation overnight. Collected flowthrough and every wash step by centrifuging 5 min 1000 xg. The beads were then washed as was done in section 3.2.7 but with 50 ml of buffers instead of 25 ml. Additionally, washed the beads with buffer D to get rid of $MgCl_2$ and ATP. Then 15 ml buffer D with 1 mM DTT was added and 150 µl protease per tube to perform on column cleavage, incubated 3h while rotating in +4°C. Centrifuged the tubes and collected a sample for gel to see if cleavage occurred. Also had collected a sample from each step to run on the same gel. No cleavage was detected so let the cleavage go on further overnight rotating in +4°C. Transferred the beads into two separate columns and collected the cleaved product through them, then

added 10 ml buffer D with 500 mM KCl and 1 mM DTT and collected the high salt wash through column. Then added 10 ml buffer D containing 15 mM glutathione, incubated for 10 min and collected the first elution (10 ml) from both. These elutions were kept separate until the gel is run. Then repeated the previous step to collect elution 2, and then added 5 ml and collected after 10 min incubation. Samples were collected from all these steps and run in the gel. According to the results of the gel 5 of the collected elutions were combined, buffer exchanged and concentrated from 40 ml to 5 ml. To determine the concentration of the protein sample run it on gel with different concentrations of standard BSA. To cleave the GST from IRE1 left it overnight with protease. The protein got precipitated during the cleavage and to see if the protein was completely lost, run a gel with samples from the mixture, the pellet and soup of the precipitated product.

3.2.14 Cleaving yIRE1-1475

From the AIM purification 4 tubes of frozen yIRE1-GST-1475 was placed on ice to melt for cleaving off the GST. Added protease for overnight cleavage, took samples from before and after cleavage. Added glutathione beads to capture free GST 2h rotating. Then centrifuged it at 1000 xg 5 min and collected the elute. Added some buffer D to the beads to collect more protein elution. Run samples on gel and aliquoted the GST-free IRE1-1475.

3.2.15 mRNA cleavage reaction

Prepared buffer E (20 mM HEPES (pH 7.5), 10 mM magnesium acetate, 50 mM potassium acetate, 70 mM NaCl, 1 mM DTT) with 2 mM ADP. For the 20 µl reaction 1.5 µM protein (with and without GST) was used and 1 µl of HAC1 mRNA was diluted with 9 µl DEPC-water. To see if the refolded mRNA binds to the IRE1 better, heated some of the mRNA to +90°C for 1 min and took the heating block to cool in its own. The cleavage reaction was prepared with both heated and unheated mRNA in separate tubes with protein and incubated 3h on +30°C and stopped with 10 µl RNA loading dye containing 0.2% SDS. The tubes were then heated to +80°C for 3 min and immediately placed on ice and loaded to RNA gels and run 180V. Used SYBR GOLD dye for 30 min covered with aluminum foil and imaged gel.

3.3 Human IRE1

The human construct was a kind gift from Genentech (Harnoss et al., 2019). The construct has a His6-tag in the beginning of the IRE1 construct. Transformation of this construct into DH10Bac (Thermo Scientific™) cells were attempted multiple times, with no results. Many factors and details during tranformation were changed, such as the amount of the plasmid, fresh antibiotics, different cell batch from other lab, fresh cells, controls etc. but it did not affect the results. There were always colonies, but all were colonies were small, and it was impossible to select the correct white colonies, therefore it was best to reclone the hIRE1 insert into another vector and start anew.

3.3.1 Preparing genetic info for bacmid

Designed primers to clone the hIRE1 and performed 100 µl PCR reaction, where the annealing temperature was +60°C. Of the PCR product 5 µl was run on 1% agarose gel (Standard Agarose – Type LE, BioNordika) with purple loading dye (New England BioLabs Inc.). Since there were multiple bands in the gel decided to redo the reaction with a bit less of genomic template and run the agarose gel. Decided to go with both PCR products and performed digestion to them and to pfastbac1 vector with restriction enzymes SalI and KpnI (New England BioLabs Inc.). Both were left to incubate in +37°C overnight. Next, loaded the samples into agarose gel and cut the correct bands from it in UV room. Purified the gel pieces according to the PCR and gel purification kit (MACHEREY-NAGEL GmbH & Co) with final elute volume of 20 µl. Measured the absorbance of them A_{260} with nanodrop spectrophotometer (vector = 40.1 ng/µl, PCR1 = 16.6 ng/µl, PCR2 = 7.3 ng/µl). Further used the PCR1 product to do the ligation in 1:2 and 1:3 ratio, mixed the ligation solution well and placed for 2h in 16-18°C. Then transformed DH5α *E. coli* cell with 5 µl of the ligation products and incubated on ice for 15 min, then plated them on LB-plates with ampicillin. Picked 5 colonies from the 1:2 ratio LB-plate into 5 µl of MQ-water and performed colony PCR with them, run the PCR product in the agarose gel. The rest of the MQ-colony solution was grown overnight in +37°C with 500 µl of LB-media and ampicillin. The overnight cultures were harvested and then purified the plasmid according to plasmid purification kit (MACHEREY-NAGEL GmbH & Co), with 30 µl of final plasmid elution. Then digested some of the samples with SalI and KpnI 1h in +37°C and run agarose gel with both undigested and digested samples. From the gel it was clear that not all samples were successful and continued with

only 3. Then thawed 3 DH10Bac *E. coli* cells and added 1 µl of the plasmids to the cells. Incubated on ice for 15 min, then added 500 µl of LB-media left it shaking in +37°C overnight and plated next day on LB-plates with kanamycin, gentamycin, tetracycline, X-gal and IPTG. The plates were covered and placed in +37°C. In a few days blue and white colonies could be detected from all of the plates and therefore restreaked 4 white colonies of each plate into new LB-plates with the same antibiotics as previously and let them grow +37°C. Grew 4 ml LB-media liquid cultures from the restreaked white colonies with kanamycin, gentamycin and tetracycline shaking +37°C overnight. The next day harvested the cells at 4000 rpm 5 min and removed the supernatant. Resuspended the pellets with the same resuspension buffer from plasmid purification kit and the added 250 µl of solution 2 (2% SDS, 200 mM NaOH) incubated 15 min on ice then added solution 3 (10% Acetic acid, 3M KAc) and gently shifted the tubes a bit and let it incubate on ice 15 min. Then centrifuged them at 14.8 rpm 15 min on the benchtop centrifuge, after which added the supernatant to new tubes with 800 µl cold isopropanol and incubated on ice for 15 min. Then centrifuged for 15 min and removed the supernatant, washed the pellet twice with 70% ethanol and centrifuged at 14.8 rpm 5 min. Then let the ethanol dry out from the tubes and added 40 µl of MQ-water. Run PCR with these bacmids after diluting some of it 10 times and checked by running the agarose gel to make sure there is bacmid (4000bp). Froze the suitable bacmids according to gel and sequencing for later use in -20°C.

3.3.2 Sf21 insect cell growth and transfection

The insect cells Sf21 were grown in SFX media (HyClone, Cytiva) at +27°C shaking 120 rpm. The cells were counted (TC20™ Automated Cell Counter, Bio-Rad Oy) and grown until the live cells count was around 80-95%. The cells were diluted to 0.5 x 10^6 cells/ml. The bacmid was prepared by adding 100 µl of filtered PBS and heated for 10 min at +70°C to avoid any contamination. After the bacmid cooled to room temperature and 5 µl of transfection reagent (Transporter 5™ Transfection Reagent 0.1 micron sterile-filtered, Polysciences, Inc.) was added and incubated 15 min. The cells were placed in a 6 well-plate, 4 ml cells/well. From these 6 wells, four wells were transfected with the bacmid. This would be the V0 generation. Let the cells grow while shaking 190 rpm at +27°C for four days. Checked the condition of the cells and found that the cells were dying as it was supposed to. Collected these cells and used them to infect fresh insect cells in two glass flasks (V1), let these grow shaking in +27°C while maintaining fresh cells to be infected if needed.

3.3.3 Expression of human IRE1

Took 1 ml samples from both V1 cultures and centrifuged them 14 rpm 5 min to collect the pellet. Used 50 µl 1xPBS and 5 µl DNase I mixture for resuspending the pellet, incubated 10 min in room temperature and freeze thawed them 4 times. Then run the gel with these samples and performed western blotting with the gel against anti-His as a primary antibody and secondary antibody anti-mouse green. Imaged the gel.

4. Results

4.1 Expression and purification of yeast IRE1

The growth of the *E. coli* BL21 RIPL cells was successful, both constructs of yeast IRE1 had plenty of colonies with a slight difference in size. The reason for this might be that the expression of one of the constructs is a bit toxic to the cell. The small-scale expression test of these both constructs was not successful (Figure 3A), therefore the expression was repeated to find the optimal IPTG concentration. In the first expression the concentration was 0.4 mM while in the second attempt we tried 1 mM of IPTG. Cells expressing yIRE1-1475 were unable to grow in liquid culture likely due to toxic expression (Table 1), thus yIRE1-1477 expression was carried out further. We also tested the effect of the temperature on cell growth and expression of protein. The 10 ml expression was divided into two tubes which were kept at different temperatures after inducing (20°C and 37°C). The induction time was increased from 2h to 4h. The expression increased clearly after inducing with 1 mM IPTG (Figure 3B). Similarly, the expression was better at 20°C than at 37°C.

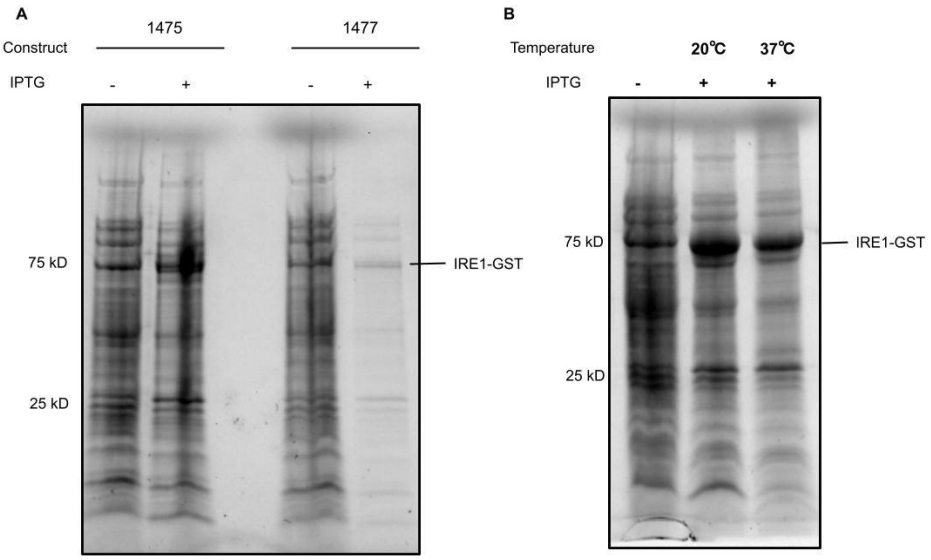

Figure 3 – Small scale expression test. (A) Small scale expression test of two yeast IRE1-GST constructs (1475 and 1477), before and after adding 0.4 mM IPTG. (B) The yeast IRE1 construct 1477 expression before and after 1 mM IPTG and 4h at two different temperatures (20°C and 30°C).

Table 1 – The measured OD_{600} from 1st and 2nd expression

1st expression		2nd expression			
Time	OD_{600} of 1477	Time	OD_{600} of 1475	Time	OD_{600} of 1477
1h	0.0474	1h	0.0267	3h	0.1384
2h	0.0626	2h	0.0276	3.5h	0.2423
3h	0.1198	3h	0.0285	4h	0.3333
4h	0.2859	4h	0.0292	5h	0.5261

The expression of construct 1477 was scaled up to 6l culture and because of the large amounts of this expression, a bit less antibiotics were used than before (see section 3.2.3), instead of 1 mM IPTG we used 0.8 mM IPTG. Nevertheless, the 6l expression was still successful as can be seen in figure 4A.

To test whether plasmid is firmly retained in the cells expressing yIRE1-1475, the growth was done overnight with carbenicillin instead of ampicillin and then diluted to OD_{600} of 0.2. This was then further grown until 0.5 and induced with 0.8 mM IPTG (Figure 4B). The expression of this construct was considerably lower than the 1477 construct. We also test auto induction media for the expression of 1475 construct.

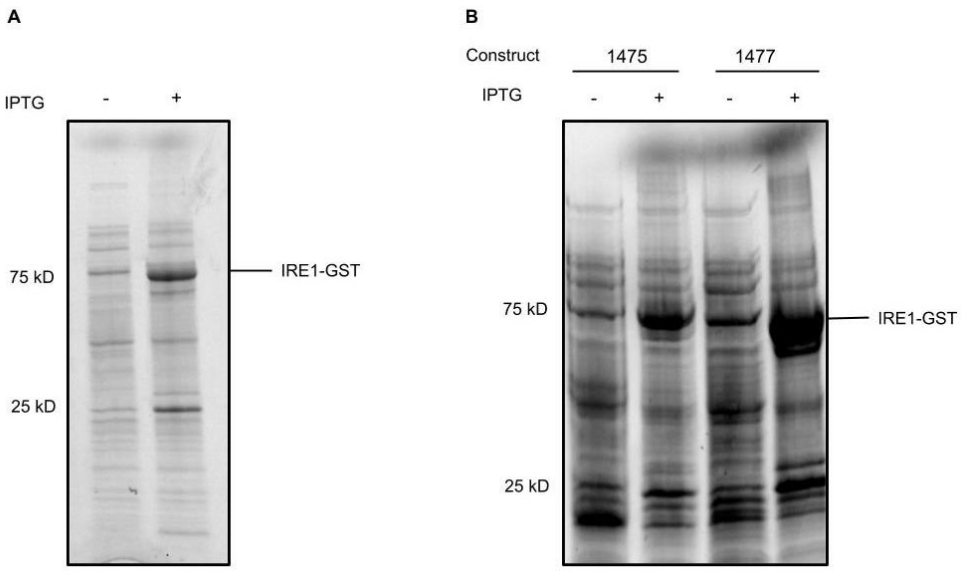

Figure 4 – Large scale expression results of construct 1475 and 1477. Expression results of constructs 1475 and 1477 before and after inducing with IPTG. (A) Expression of 6l of yeast construct 1477 with 0.8 mM IPTG (B) Construct 1475 was expressed using carbenicillin instead of ampicillin. The 1l expression result of construct 1477 with 0.8 mM IPTG.

Purification of yIRE1-1477 was done with 1l cell pellet first, that was kept separate from the 5l. The binding capacity of the glutathione sepharose beads is not as high as expected (5 mg GST per ml beads) there was still plenty of protein in the flowthrough after 1h binding. SDS-PAGE gel analysis shows undetectable amount of protein in wash 1 (Figure 5). The impurities were removed by gently washing the beads with wash buffers through a column and no impurities were visible in gel image after wash 1.

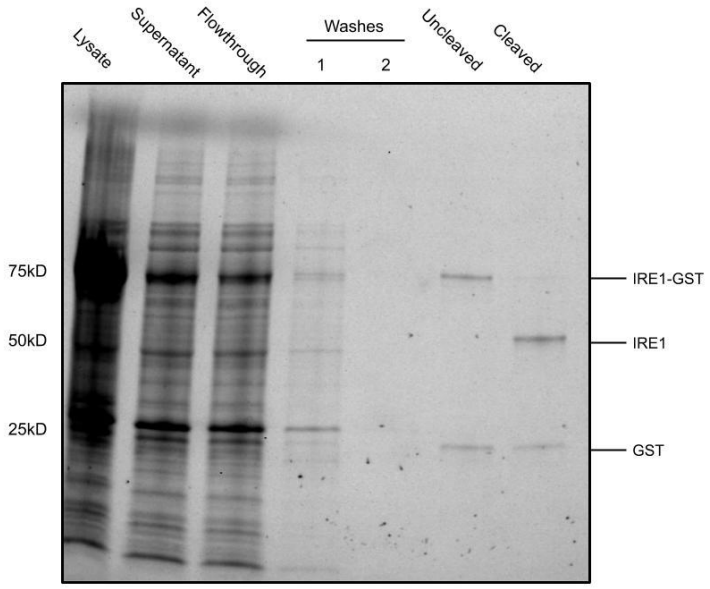

Figure 5 – SDS-PAGE analysis of purification. SDS-PAGE gel analysis of various fractions of collected washing steps during purification of yIRE1-GST construct 1477. Lysate is from lysing cells with emulsiflex. Supernatant is after centrifugation that was then added to the beads for binding. The flowthrough is the collected sample after binding to the beads. Washes through the column got rid of impurities already after first wash. The elutes were combined as uncleaved sample and cleaved sample was eluted through column to get rid of free-GST.

Attempts to remove GST tag from the target protein using on-column cleavage with 3C protease did not seem to work, which is why, the release of the bound protein with glutathione was needed. After collecting elutions with glutathione the measured A_{280} and the gel image of the elutions indicated that there is protein even though the absorbance values

were low, likely due to fever number of aromatic amino acids and the shielding effects of yIRE1 oligomerization on these residues (Figure 6 and Table 2).

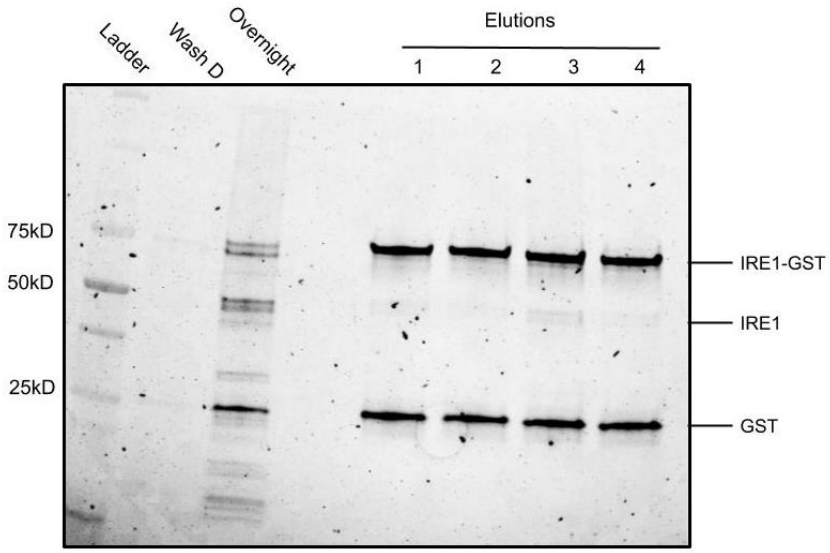

Figure 6 – SDS-PAGE gel of the purified yIRE1-GST. SDS-PAGE gel of the purified protein with All blue ladder (Bio-Rad Cat #161-0373). Most of the protein was in elution even though A_{280} did not indicate so (Table 2). Wash D is from washing the beads from the overnight rotation with protease 3C. Overnight sample is the flow through from the previously mentioned overnight rotation. Elutions from beads with glutathione contains both yIRE1-GST and free GST, with a faint band of GST free yIRE1.

Table 2 – The measured A_{280} from 1l purification

A_{280}	
Elution	mg/ml
1	0.089
2	0.108
3	0.178
4	0.142

Therefore, combined elutions were mixed with protease to cleave the GST from IRE1 after which beads to the sample to capture the free GST and uncleaved protein. According to the SDS-PAGE gel analysis, fraction of the protein was successfully cleaved and eluted, but some had remained in the column which was eluted with glutathione to get the yIRE1-bound GST with free GST (Figure 7). The A_{280} was measured after concentrating the samples and 20% glycerol was added to be aliquoted and stored in -80°C. The concentration was calculated with an online calculator (Protein Parameters, protparam.net/index) that calculates the extinction coefficient from the protein sequence and A_{280} values (Table 3).

The concentration of the protein with the GST is higher than without GST, yet there is enough to be used later for mRNA cleavage assay.

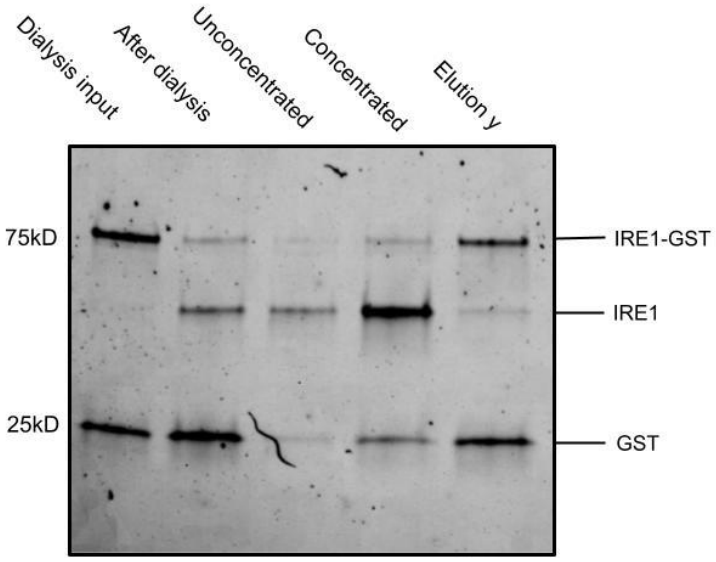

Figure 7 – SDS-PAGE gel of construct 1477 after cleaving GST. SDS-PAGE gel of construct 1477 after cleaving GST during dialysis (see dialysis input and after dialysis). That sample was then added to glutathione sepharose beads and the 5 ml elution (Unconcentrated) was collected to be concentrated. Eluted from the previous beads with glutathione elution y.

Table 3 – The concentration of purified yIRE1-1477

yIRE1-1477			
Sample	A_{280} mg/ml	Extinction coefficient	Concentration µM
IRE1	0.124	40340	3.03
IRE1-GST	0.602	69680	8.64

Since the construct yIRE1-1475 did not grow as well as construct yIRE1-1477 due to toxic expression, which is why autoinduction media (AIM) was used for expression. Both constructs were treated with same antibiotics as before with slight difference; carbenicillin for yIRE1-1475 due to low growth rate with ampicillin, which could have turned toxic to the cells and ampicillin for yIRE1-1477. The 1 ml samples from before AIM and after proved to be tricky to load into the SDS-PAGE gel, after a redo it was managed but there were no uninduced samples to run along the induced one (Figure 8). Even without the proper compare to the uninduced sample, the AIM did show expression especially with construct

yIRE1-1477, but even so the expression of construct yIRE1-1475 looked better than it has previously. This test run was done with approximately 600 ml of media, which is why it was centrifuged to keep the pellets for purification.

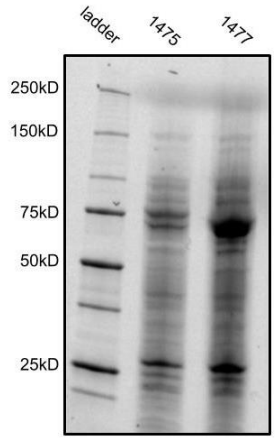

Figure 8 – The results of the autoinduction media expression. The SDS-PAGE gel image of the AIM expression with IRE1-GST samples of both constructs (1475 and 1477) after induction.

The purification process was same as previously but the binding into the beads was done overnight, there did not seem to be much of a difference as can be seen while comparing figures 5 and 9. Also the protein eluted at different time than before, construct during wash B and construct 1477 during wash B and ATP. This seemed odd, the only thing that did change was the DTT used in these washes. To test this assumption, the purification was just the same way as previously with yIRE1-GST construct 1477 with a different batch of DTT to see if that changes anything.

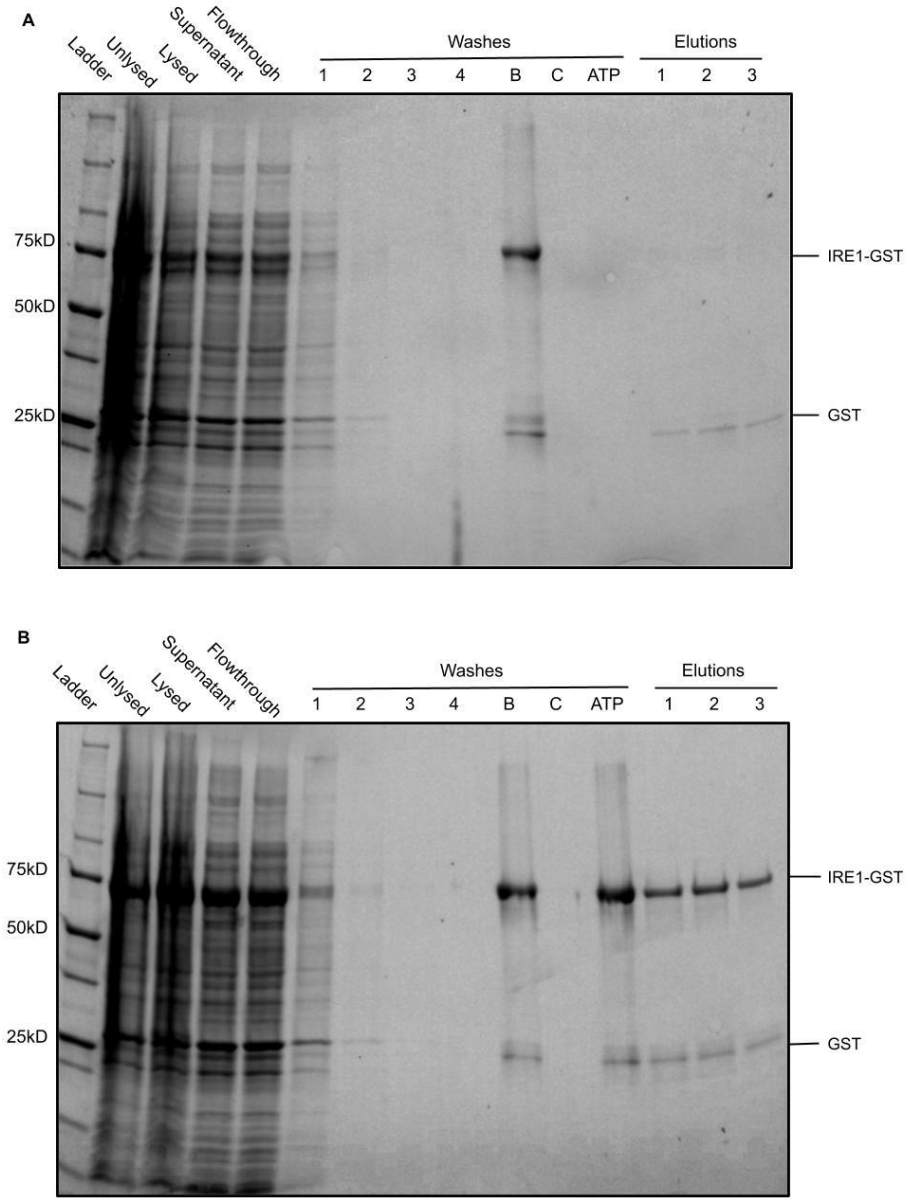

Figure 9 – Purification process of AIM expressed yIRE1-GST. SDS-PAGE gel of purification process of AIM expressed yIRE1-GST constructs 1475 and 1477. (A) Construct 1475 yIRE1-GST purification process. The protein was eluted out of the beads in wash B instead of elutions. Elutions contained only GST from previous purifications. (B) Construct 1477 yIRE1-GST purification process. The protein did elute as should but also a lot of protein came in wash B and ATP. Also, free GST most likely from previous purifications.

Luckily a different DTT was the solution to this problem, as can be seen from figure 10B and C. At the same time the bead-to-sample ratio was tested, with 15 ml (Figure 10B) and 2 ml (Figure 10C) to same ratio of beads. The gels indicate that the difference in beads to solution ratio does not make much of a difference, the binding capacity of the glutathione sepharose beads is not as great as wished for (stated binding capacity in instructions of the Sepharose

glutathione 4B beads was 5 mg GST/ml medium). If the binding capacity of the beads were greater than what it was then also the yield would be much higher.

Even though the protein got eluted at a different time, these samples were still collected and concentrated to smaller volumes (Figure 10A). According to the gel there seems to be a slight size difference between the two constructs, construct 1475 is a bit higher on the gel that the construct 1477 (Figure 10A). The measured A_{280} of the concentrated samples from AIM expression and calculated their concentration the same way as explained above (Table 4).

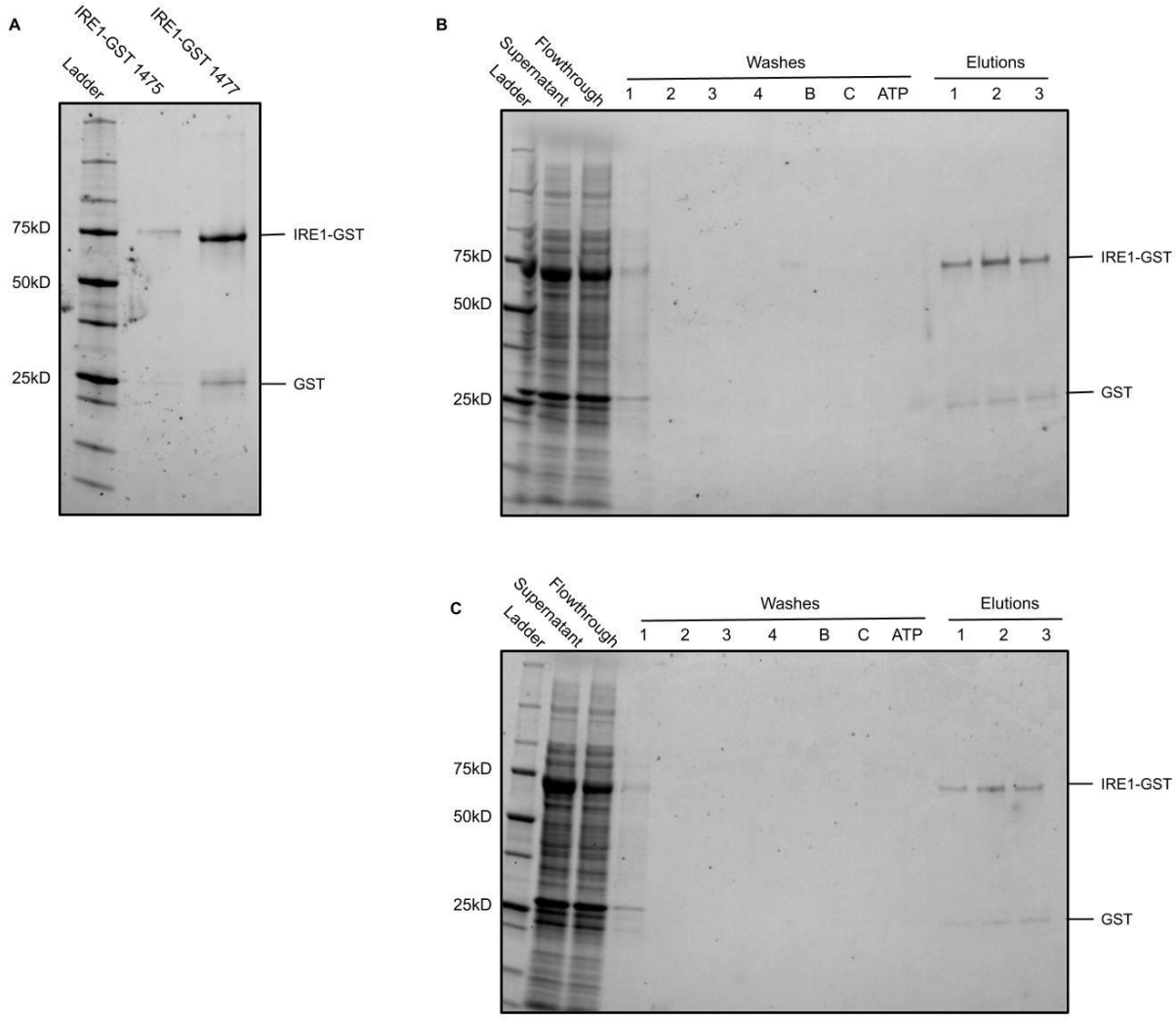

Figure 10 – The SDS-PAGE gels of AIM expressed yIRE1-GST. (A) Concentrated samples from AIM expressed protein purification. There is a slight size difference between the two constructs. (B) 15 ml reattempt protein purification with new DTT and 1.5 ml beads. (C) 2ml reattempt protein purification with new DTT and 1.5 ml beads.

Table 4 – The concentration of purified yIRE1 1475 and 1477 from AIM expression

yIRE1-GST			
Sample	A_{280} mg/ml	Extinction coefficient	Concentration µM
1475	0.068	69680	0.98
1477	0.612	69680	8.78

Both constructs were now available for mRNA cleavage assay, the assay were performed with the AIM expressed proteins with GST. The gel image was not clear, and it seemed like there was a problem with the HAC1 mRNA, it was needed to be treated with DNase. Also, the replacement of ADP with ATP could be one of the reasons why the cleavage did not happen. It could be that the yIRE1 protein requires ATP to be stable and active.

While waiting for better HAC1 quality mRNA, started purifying the 5l pellet that was stored before in -80°C to see if we could get a high enough yield for crystallization and X-ray. For the purification the binding was done overnight, and all steps were made in falcon tubes, which is the reason why the washing steps carried a lot of impurities beyond wash 1 (Figure 11). The protein got eluted a bit later than before, since the elutions were gathered though a column. And the first 2 elutions are still part of the wash ATP, due to this more elutions were gathered (8 elutions in total 5 ml each) and their A_{280} was measured. All of them were combined and concentrated to 5 ml and the A_{280} was 9.123 mg/ml. During A_{280} measurement noticed that the peak in the measuring program with the nanodrop was not in 280 but in 260. This could be the RNase domain, that is overtaking the results of the absorbance. Which indicates that the A_{280} value may not be correct. To get GST cleaved from yIRE1 added protease to it overnight and added beads to them next day to capture the free GST. After which Ni-NTA beads were added to capture the protease, when sample was collected and concentrated the sample from 17 ml to 3 ml and the measured A_{280} was 6.655 mg/ml.

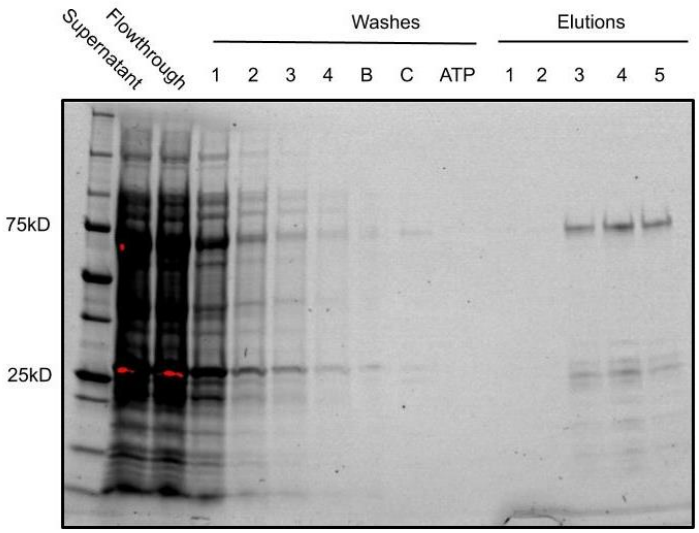

Figure 11 – Purification steps of construct 1477 5l pellet. Impurities were carried beyond wash 1 due to the use of a falcon tube instead of a column, but these impurities were completely washed away in wash ATP and the protein got eluted in elutions 3, 4 and 5.

For crystallization the protein should be as pure as possible, which is why the purified protein was injected into the gel filtration column. During filtration the 4 peaks were observed, and their corresponding fractions collected. The assumption was that one of the two major peaks would have been our protein, but when tested the fractions on gel there was nothing visible. We first thought that maybe it is too diluted, and concentrated the fractions, but still in gel nothing was visible. Therefore, other peaks were investigated, but at the end it looked like the protein was lost somewhere in the process. Which is why we repeated everything with 12l expression, everything went well until the cleavage of the GST from yIRE1. The protein precipitated and was unsalvable. Unfortunately, crystallization was no longer an option in this project.

In the -80°C storage were previously purified proteins stored for activity assay, which was the next aim. There were all proteins except yIRE1-75 without GST, which is why some of the stored tubes were combined to be cleaved with protease. Then used thoroughly cleaned beads to capture free GST. Collected the elution from the beads and run a gel with the sample, but unfortunately it seemed like the sample got contaminated with free GST and uncleaved protein. Which raised a question of if the beads need a special cleaning procedure, to get them clean. In the end we had yIRE1-75 sample mixture with GST bound and without (Figure 12A).

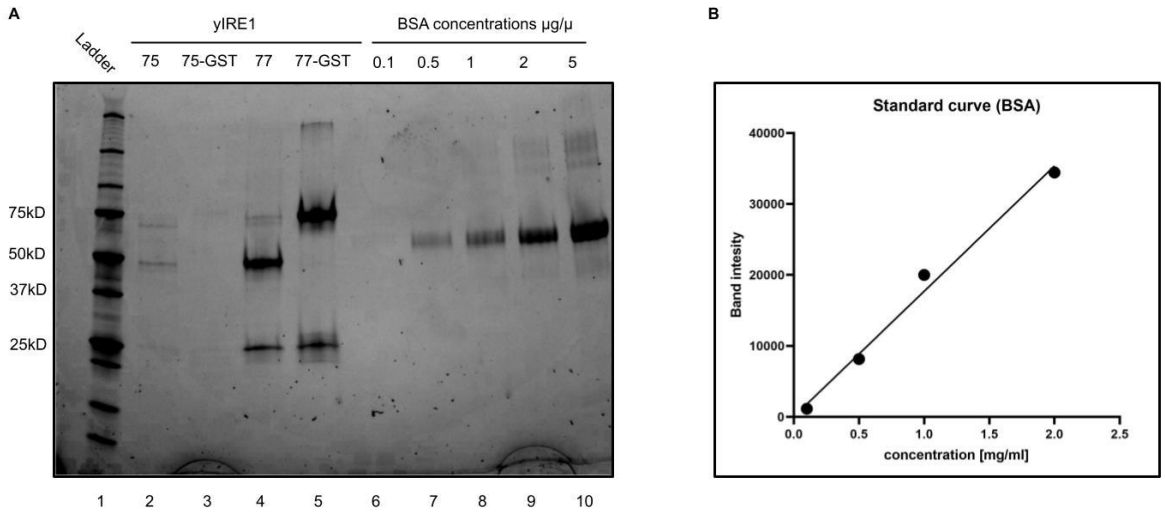

Figure 12 – Quantification of purified yIRE1 using SDS-PAGE Gel BSA standards. (A) The concentrations of yIRE1 proteins were calculated from the gel by comparing the band intensity to the known concentrations of BSA bands with different concentrations. (B) The standard curve of BSA band intensity and the concentration.

Table 5 – The concentration yIRE1 1475 and 1477 based on the BSA standard curve

BSA standard curve concentrations		
Sample	mg/ml	Concentration µM
1475	0.275	5.03
1475-GST	0.225	2.83
1477	1.775	32.47
1477-GST	1.575	19.83

Once HAC1 mRNA was available and the concentration of the proteins recalculated with BSA-standard curve (Figure 12 and Table 5) the mRNA cleavage assay was attempted several times. Mostly there was not much change in reaction, the concentrations of protein were at first 1µM then it was increased to 1.5µM. Also, the mRNA was first diluted 10-fold, since it had a strong signal on agarose gel, but then used the undiluted version. There was an assumption that maybe the mRNA does not bind well in a coiled from and needs to be refolded by heating it to +90°C and letting it cool down in room temperature. In gel we tried to both mRNAs, unheated and heated. Still there was not much to be seen in the gel. At the very last attempt everything was redone and the only changing factor was the dyeing of the gel, which was done with fresh SYBR Gold dye and covered with aluminum foil for 30 min when we usually were using a over a half year old dye with no covering and only for 10 min.

This was the change that needed to be done, there was finally something visible in the gel (Figure 13). And in the gel, it was detectable that most of the proteins were active, there were smaller mRNA products in protein containing samples as in the control. In addition, the gel showed that the mRNA did not need to be refolded for the cleavage to happen.

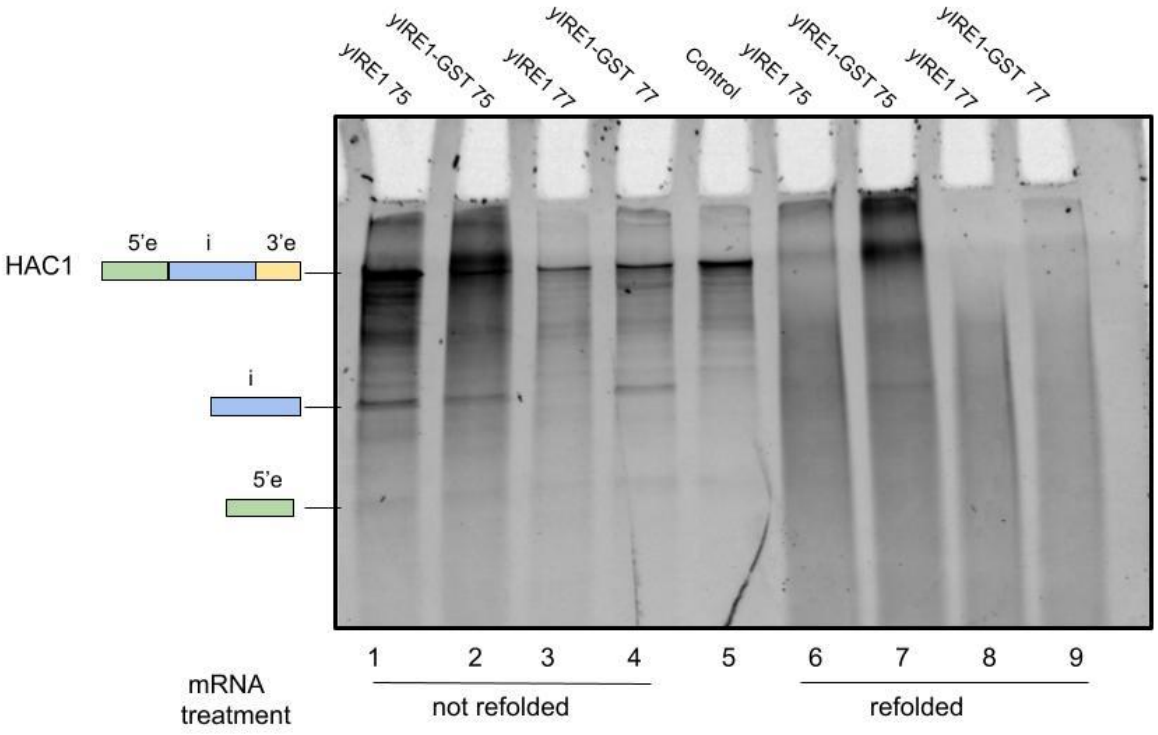

Figure 13 - mRNA cleavage assay of isolated IRE1 protein. Cleavage happened with most proteins, but not with GST free construct 1477. Refolding the mRNA with heating did not result in cleavages except with construct 1475 with GST. Also, the smallest exon from 3'-end seems to have run out of the gel, due to its size.

4.2 Results of human IRE1

The human IRE1 was a challenge, lots of time was used for blue and white screening, which is a quick technique to identify the bacteria that has the desired sequence successfully inserted by using the activity of an enzyme (β-galactosidase) in *E.coli*. β-galactosidase is an enzyme that hydrolyses X-gal which will turn the cell blue. Cells containing the insert will turn white, since the activity of the β-galactosidase has been disrupted, and cells that do not have the insert will be blue. We transformed self-prepared DH10Bac competent *E.coli* cells carrying the Bac-to-Bac bacmid with our plasmid, positive bacterial colonies with the desired

sequence should be white in color and the negative blue. There seemed to be a problem with the construct that we had received. According to my results that is based on multiple tries of transformation, with fresh cells, different cells from neighboring lab, fresh plasmid, different amounts of cells and plasmid plated or added, fresh antibiotics, and with using controls. There just seemed to be a difficulty to identify the transformed cell from non-transformed ones. Which could also indicate that there is a problem with the antibiotic resistances in the Tn5 site. Therefore, the human IRE1 sequence was recloned into pFastBac1 vector plasmid, which then produced positive white colonies along the blue ones. After a successful cloning transformed DH10Bac *E.coli* cells with the new plasmid, and successfully got blue and white colonies. From some of these individual colonies liquid cultures were grown and the bacmid was isolated and stored for later transfection into Sf21 insect cells.

The first batch of Sf21 insect cells were not growing as supposed to and there seemed to be some issues with the maintenance. Which is why new batch of cells were thawed and were maintained to stay in 0.5 million cells per ml for the transfection. During this maintenance we noticed that the dye that was used to dye the cells to be counted by the cell counter was contaminated and probably old, there was no clear indication when the bottle was opened.

When the cells were ready to be transfected, the plasmid was prepared for transfection with transfection reagent. The virus stock was ready in 5 days according to the protocol we followed (protocol was prepared by a lab member) since the cells were dying according to the cell counter machine as it was supposed to. Some of the virus was then stored and some used to infect a fresh insect cell culture, but it seemed like the cells were doing fine and were not dying as supposed. To test this, we took some cells to see if there was expression of hIRE1 detectable by running a gel and performing western blotting. In western blot image nothing could be seen, even though we had a positive control which indicates that the virus infection was not successful.

5. Discussion

5.1 Differences between the yeast IRE1 constructs

While expressing the both constructs, we had a hunch that the one that did not express well would be the active construct and the other would be inactive. Therefore, when we got the results of mRNA cleavage assay the 1477 construct alone without GST did not cleave the mRNA as the other construct did. This indicated that the easy expression of 1477 was due to its inactivity, while the active protein seemed to be difficult to express due to its toxicity to the cell. It would have made a difference to know which is which between these two constructs, but every attempt to sequence these constructs resulted in failure, since the beginning of the IRE1 sequence was always badly sequenced. The only differences we knew about these constructs was in the beginning of the IRE1 sequence, the other is longer that the other (see APPENDIX 1). During AIM expression we got both constructs next to each other on a gel, which then indicated that the construct 1477 might be a bit smaller than construct 1475 (Figure 8). But when all of them were in the same gel with BSA (Figure 12) there does not seem to be a difference in size between the GST free constructs 1477 and 1475, but with the GST-tagged construct there is a difference in size. This raises the question of why the GST-tagged proteins differ in size, and the non-tagged not.

Both constructs were a kind gift from Peter Walters lab, and according to their literature they had a few variants of the cytosolic IRE1 (Korennykh et al., 2008). All contained the RNase domain and the kinase domain, but the difference is in amino terminus of the IRE1 proteins. It was extended by 24, 32 or 120 amino acids, which are from the linker domain, that binds the cytosolic part to the transmembrane part on the ER membrane. We only know that the construct we have are on that is just the kinase and RNase domain and the other one is extended by 32 amino acids. But we do not know which of the constructs is which, and the sequencing was not successful in the desired extension area after multiple tries it was almost impossible to know which is the extended one and which is not. This needs to be investigated further by other lab members.

5.2 Purification process and removing GST

During the purification everything went smoothly, with no major problems and even the number of washing steps made sense when the washing was done in a 50ml Falcon tubes. The purification on itself caused no problems, except when the proteins eluted at the wrong

wash (Figure 9). At that time the buffers had been made again, which raises the issue of being absolute sure of the concentrations of the reagents that you use and that there is a chance of human error, maybe the added amounts were not right to begin with.

The cleavage of the GST-tag was not as easy and straightforward as it sounded on the article that we followed (Nock et al., 2001). In the article the cleavage was performed on column while the protein was bound to the beads with protease, but in our case, nothing got cleaved. It did not matter how much protease we used, how long or what kind of protease. We had both commercial and self-prepared, but the results were same with both, no cleavage. Therefore, we had to elute the protein from the beads with glutathione and cleave off the GST without the beads. Then after cleavage the GST had to be rebound to the beads, while also the noncleaved protein got bound, to get the GST-free IRE1. This procedure was effective but took a lot of time and sometimes even contaminated our samples with previously purified IRE1 proteins, even though the beads had been cleaned thoroughly with water and buffers. There is a possibility that the cleavage site of IRE1-GST when bound to the beads in the column was not accessible but when it was unbound in a solution the site was exposed to be cleaved.

The separation of IRE1 from GST has not always worked, we still are not quite sure at what step we had lost our protein that was meant to be gel filtrated. It was supposed to be injected to the gel filtration system, but according to the collected peak fractions gels the protein was not found. The question is, did it elute before we started collecting fractions, was it still bound to the beads, or did it stay in the column. Even the reattempt of this whole process was unsuccessful and lead to precipitated protein. It was stated by Korennykh et al. 2018 that the protein needs during fast protein liquid chromatography (FPLC) at least 300mM of NaCl to prevent its aggregation. Not only did we learn this knowledge late since it was not stated to our knowledge by Nock et. al. 2001, but also thought that maybe other salts would have had the same effect. To sum up everything that has been stated this far, more thorough reading of the newer publications might have gotten us better results.

5.3 Activity of yIRE1 in mRNA cleavage assay

The yeast IRE1 is supposed to cleave an intron from between two exons in HAC1 mRNA, this was successfully tested in the last RNA gel. The assay was repeated multiple times without knowing what was wrong, often the mRNA quality was questioned and the procedure of the whole process. Also, the performance of during the assay was questioned, until the

real issue was found. The usage of SYBR Gold dye was not commonly known in our group, which is why we did not know that the dye would expire so quickly. Once this was found out the results of the RNA gels could finally be interpreted.
The construct that was not expressed easily was active, and the other only with GST-tag. This could mean that the expression of the protein is difficult because it if functional which could be the reason why it was toxic to the cells to express. Additionally, the easily expressed protein was active with the GST-tag, this could mean that the GST somehow stabilizes the protein to be active.

5.4 Human IRE1

The human construct was not what we expected it to be in the beginning, it took a while to notice that it was not working, not due to human error during laboratory work or in methods, but due to the construct itself. Once the IRE1 sequence was cloned to our own vectors, the transformation of the cells was successful and white colonies detected. But at the time this was noticed, and results achieved, precious time had already been lost and the remaining time was not sufficient to get actual results.
We made and attempt to get any kind of expression with infecting insect cells, but it did not work at the first time around, and the maintenance of the insect cells proved to be difficult. Only later did we learn that the cell counter was malfunctioning and probably not even giving right counts, we relied on that counter too much. Sometimes we were a bit suspicious about the dye we used and checked if the counter gave a realistic result by looking at the cells with a microscope. With all this it did not cross our minds to suspect the cell counter itself, also there is a possibility that the counter was functioning during our experiments and started malfunctioning afterwards.

6. Conclusion

The yeast IRE1 protein was easy to express and with a bit of extra work purified. The activity of the yeast protein was proven, and it also showed that the protein could be active with the GST-tag. This actively demonstrated that the removal of the GST is not entirely necessary and there is a possibility to get a crystal structure of this protein with the GST. Since most of the trouble was in removal of the tag and not expression. Even the toxic protein was expressed in some amount with the AIM, we only would have needed to raise the expressed volume and there would have been enough for crystallization. The human IRE1 in other hand just needed more time to be expressed properly, and later hIRE1 was expressed in a high yield, purified and its activity proved by other members of the lab, which means that the recloned construct has worked. There is a high chance the future research of IRE1 would lead to the ability to increase the protein folding capacity of the cell with the help of small molecules. This would be the next step to find solutions to diseases where mutations cause protein misfolding.

7. Acknowledgements

I wish to thank Ville Paavilainen for all the help and advice he has given me and for giving me the chance to work in this project. I thank Shahid Rehan for his patience, support, advices, and teachings during laboratory work. I highly appreciate every lesson you gave me. Also, I wish to thank the following people for helping, giving advice and instructions when I needed; Anja Paatero, Juho Kellosalo, Vanessa Bensch, Katja Rosti, Sari Tynkkynen, Marina Green and Dale Tranter. I wish to thank professors Pia Siljander and Kari Keinänen for support and advice.

Finally, I will always thank Saed Adawi for his support and encouragements, it was and is needed and very much appreciated.

References

Adams, C.J., Kopp, M.C., Larburu, N., Nowak, P.R., and Ali, M.M.U. (2019). Structure and Molecular Mechanism of ER Stress Signaling by the Unfolded Protein Response Signal Activator IRE1. Frontiers in Molecular Biosciences *6,* 11.

Ali, M.M., Bagratuni, T., Davenport, E.L., Nowak, P.R., Silva-Santisteban, C., Hardcastle, A., Mcandrews, C., Rowlands, M.G., Morgan, G.J., Aherne, W.*, et al.* (2011). Structure of the Ire1 autophosphorylation complex and implications for the unfolded protein response derivative works without specific permission. The EMBO Journal *30,* 1-12.

Amin-Wetzel, N., Saunders, R.A., Kamphuis, M.J., Rato, C., Preissler, S., Harding, H.P., and Ron, D. (2017). A J-Protein Co-chaperone Recruits BiP to Monomerize IRE1 and Repress the Unfolded Protein Response. Cell (Cambridge) *171,* 1625-1637.e13.

Cabral-Miranda, F., and Hetz, C. (2018). ER Stress and neurodegenerative disease: a cause or effect relationship? Current Topics in Microbiology and Immunology *414,* 131-157.

Calfon, M., Zeng, H., Urano, F., Till, J.H., Hubbard, S.R., Harding, H.P., Clark, S.G., and Ron, D. (2002). IRE1 couples endoplasmic reticulum load to secretory capacity by processing the XBP-1 mRNA. Nature (London) *415,* 92-96.

Carrara, M., Prischi, F., Nowak, P.R., Kopp, M.C., and Ali, M.M. (2015). Noncanonical binding of BiP ATPase domain to Ire1 and Perk is dissociated by unfolded protein CH1 to initiate ER stress signaling. eLife *4,* 1-16.

Concha, N.O., Smallwood, A., Bonnette, W., Totoritis, R., Zhang, G., Federowicz, K., Yang, J., Qi, H., Chen, S., Campobasso, N.*, et al.* (2015). Long-Range Inhibitor-Induced Conformational Regulation of Human IRE1α Endoribonuclease Activity. Molecular Pharmacology *88,* 1011-1023.

Cox, J.S., Shamu, C.E., and Walter, P. (1993). Transcriptional induction of genes encoding endoplasmic reticulum resident proteins requires a transmembrane protein kinase. Cell (Cambridge) *73,* 1197-1206.

Cox, J.S., and Walter, P. (1996). A Novel Mechanism for Regulating Activity of a Transcription Factor That Controls the Unfolded Protein Response. Cell (Cambridge) *87,* 391-404.

Credle, J.J., Finer-Moore, J.S., Papa, F.R., Stroud, R.M., and Walter, P. (2005). On the Mechanism of Sensing Unfolded Protein in the Endoplasmic Reticulum. Proceedings of the National Academy of Sciences - PNAS *102,* 18773-18784.

Crowley, K.S., Reinhart, G.D., and Johnson, A.E. (1993). The signal sequence moves through a ribosomal tunnel into a noncytoplasmic aqueous environment at the ER membrane early in translocation. Cell (Cambridge) *73,* 1101-1115.

Ellis, J. (1987). Proteins as molecular chaperones. Nature (London) *328,* 378-379.

Gemmer, M., and Förster, F. (2020). A clearer picture of the ER translocon complex. Journal of Cell Science *133,* jcs231340.

Grabski, A., Drott, D., and Mehler, M. (2005). The Overnight Express Autoinduction System: High-density cell growth and protein expression while you sleep. Nature Methods *2,* 233-235.

Guo, Z., Hong, S.Y., Wang, J., Rehan, S., Liu, W., Peng, H., Das, M., Li, W., Bhat, S., Peiffer, B.*, et al.* (2019). Rapamycin-inspired macrocycles with new target specificity. Nature Chemistry *11,* 254-263.

Hampton, R.Y. (2002). ER-associated degradation in protein quality control and cellular regulation. Current Opinion in Cell Biology *14,* 476-482.

Harding, H.P., Zhang, Y., and Ron, D. (1999). Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. Nature (London) *397,* 271-274.

Harnoss, J.M., Le Thomas, A., Shemorry, A., Marsters, S.A., Lawrence, D.A., Lu, M., Chen, Y.A., Qing, J., Totpal, K., Kan, D.*, et al.* (2019). Disruption of IRE1α through its kinase domain attenuates multiple myeloma. Proceedings of the National Academy of Sciences of the United States of America *116,* 16420-16429.

Hartl, F.U. (1996). Molecular chaperones in cellular protein folding. Nature (London) *381,* 571-580.

Haze, K., Yoshida, H., Yanagi, H., Yura, T., and Mori, K. (1999). Mammalian Transcription Factor ATF6 Is Synthesized as a Transmembrane Protein and Activated by Proteolysis in Response to Endoplasmic Reticulum Stress. Molecular Biology of the Cell *10,* 3787-3799.

Hebert, D.N., and Molinari, M. (2007). In and Out of the ER: Protein Folding, Quality Control, Degradation, and Related Human Diseases. Physiological Reviews *87,* 1377-1408.

Hetz, C. (2012). The unfolded protein response: controlling cell fate decisions under ER stress and beyond. Nature Reviews. Molecular Cell Biology *13,* 89-102.

Hetz, C., Zhang, K., and Kaufman, R.J. (2020). Mechanisms, regulation and functions of the unfolded protein response. Nature Reviews. Molecular Cell Biology *21,* 421-438.

Hollien, J., and Weissman, J.S. (2006). Decay of Endoplasmic Reticulum-Localized mRNAs During the Unfolded Protein Response. Science (American Association for the Advancement of Science) *313,* 104-107.

Johnson, A.E., and van Waes, M.A. (1999). THE TRANSLOCON: A Dynamic Gateway at the ER Membrane. Annual Review of Cell and Developmental Biology *15,* 799-842.

Joshi, A., Newbatt, Y., McAndrew, P.C., Stubbs, M., Burke, R., Richards, M.W., Bhatia, C., Caldwell, J.J., McHardy, T., Collins, I., and Bayliss, R. (2015). Molecular mechanisms of

human IRE1 activation through dimerization and ligand binding. Oncotarget *6,* 13019-13035.

Kimata, Y., Oikawa, D., Shimizu, Y., Ishiwata-Kimata, Y., and Kohno, K. (2004). A Role for BiP as an Adjustor for the Endoplasmic Reticulum Stress-Sensing Protein Ire1. The Journal of Cell Biology *167,* 445-456.

Kopp, M.C., Nowak, P.R., Larburu, N., Adams, C.J., and Ali, M.M. (2018). In vitro FRET analysis of IRE1 and BiP association and dissociation upon endoplasmic reticulum stress. eLife *7,* 1-13.

Korennykh, A.V., Egea, P.F., Korostelev, A.A., Finer-Moore, J., Zhang, C., Shokat, K.M., Stroud, R.M., and Walter, P. (2008). The unfolded protein response signals through high-order assembly of Ire1. Nature (London) *457,* 687-693.

Korennykh, A.V., Korostelev, A.A., Egea, P.F., Finer-Moore, J., Stroud, R.M., Zhang, C., Shokat, K.M., and Walter, P. (2011). Structural and functional basis for RNA cleavage by Ire1. BMC Biology *9,* 47.

Kozak, M. (2002). Pushing the limits of the scanning mechanism for initiation of translation. Gene *299,* 1-34.

Lee, K.P.K., Dey, M., Neculai, D., Cao, C., Dever, T.E., and Sicheri, F. (2008). Structure of the Dual Enzyme Ire1 Reveals the Basis for Catalysis and Regulation in Nonconventional RNA Splicing. Cell (Cambridge) *132,* 89-100.

Litwack, G. (2018). Chapter 11 - Protein Biosynthesis. In Human Biochemistry, Litwack, Gerald ed., (Boston: Academic Press) pp. 319-336.

Mackinnon, A.L., Paavilainen, V.O., Sharma, A., Hegde, R.S., and Taunton, J. (2014). An allosteric Sec61 inhibitor traps nascent transmembrane helices at the lateral gate. eLife *3,* e01483.

Meusser, B., Hirsch, C., Jarosch, E., and Sommer, T. (2005). ERAD: the long road to destruction. Nature Cell Biology *7,* 766-772.

Nock, S., Gonzalez, T.N., Sidrauski, C., Niwa, M., and Walter, P. (2001). [1] Purification and activity assays of the catalytic domains of the kinase/endoribonuclease Ire 1p from Saccharomyces cerevisiae. In Methods in Enzymology, Elsevier Science & Technology) pp. 3-10.

Osborne, A.R., Rapoport, T.A., and van den Berg, B. (2005). Protein translocation by the Sec61/SecY channel. Annual Review of Cell and Developmental Biology *21,* 529-550.

Papa, F.R., Zhang, C., Shokat, K., and Walter, P. (2003). Bypassing a Kinase Activity with an ATP-Competitive Drug. Science *302,* 1533-1537.

Preiss, T., and W. Hentze, M. (2003). Starting the protein synthesis machine: eukaryotic translation initiation. BioEssays *25,* 1201-1211.

Prischi, F., Nowak, P.R., Carrara, M., and Ali, M.M.U. (2014). Phosphoregulation of Ire1 RNase splicing activity. Nature Communications *5,* 3554.

Ron, D., and Walter, P. (2007). Signal integration in the endoplasmic reticulum unfolded protein response. Nature Reviews. Molecular Cell Biology *8,* 519-529.

Schuck, S., Prinz, W.A., Thorn, K.S., Voss, C., and Walter, P. (2009). Membrane Expansion Alleviates Endoplasmic Reticulum Stress Independently of the Unfolded Protein Response. The Journal of Cell Biology *187,* 525-536.

Shen, J., Chen, X., Hendershot, L., and Prywes, R. (2002). ER Stress Regulation of ATF6 Localization by Dissociation of BiP/GRP78 Binding and Unmasking of Golgi Localization Signals. Developmental Cell *3,* 99-111.

Soto, C. (2003). Unfolding the role of protein misfolding in neurodegenerative diseases. Nature Reviews. Neuroscience *4,* 49-60.

Tirasophon, W., Welihinda, A.A., and Kaufman, R.J. (1998). A stress response pathway from the endoplasmic reticulum to the nucleus requires a novel bifunctional protein kinase/endoribonuclease (Ire1p) in mammalian cells. Genes & Development *12,* 1812-1824.

Vattem, K.M., and Wek, R.C. (2004). Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. Proceedings of the National Academy of Sciences - PNAS *101,* 11269-11274.

Walter, P., and Ron, D. (2011). The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation. Science (American Association for the Advancement of Science) *334,* 1081-1086.

Wang, M., and Kaufman, R.J. (2016). Protein misfolding in the endoplasmic reticulum as a conduit to human disease. Nature (London) *529,* 326-335.

Zhou, J., Liu, C.Y., Back, S.H., Clark, R.L., Peisach, D., Xu, Z., and Kaufman, R.J. (2006). The crystal structure of human IRE1 luminal domain reveals a conserved dimerization interface required for activation of the unfolded protein response. Proceedings of the National Academy of Sciences - PNAS *103,* 12343.

APPENDIX 1

The amino acid sequence of yeast IRE1-GST 1475 and 1477

MGDNIVLYYF DARGKAELIR LIFAYLGIEY TDKRFGVNGD AFVEFKNFKK EKDTPFEQVP
ILQIGDLILA QSQAIVRYLS KKYNICGESE LNEFYADMIF CGVQDIHYKF NNTNLFKQNE
TTFLNEDLPK WSGYFEKLLK KNHTNNNNDK YYFVGNNLTY ADLAVFNLYD DIETKYPSSL
KNFPLLKAHN EFISNLPNIK NYITNRKESV Y[P]EKKKRKRGSR GGKKGRKSRI
ANIPNFEQSL KNLVVSEKIL GYGSSGTVVF QGSFQGRPVA VKRMLIDFCD IALMEIKLLT
ESDDHPNVIR YYCSETTDRF LYIALELCNL NLQDLVESKN VSDENLKLQK EYNPISLLRQ
IASGVAHLHS LKIIHRDLKP QNILVSTSSR FTADQQTGAE NLRILISDFG LCKKLDSGQS
SFRTNLNNPS GTSGWRAPEL LEESNNLQCQ VETEHSSSRH TVVSSDSFYD PFTKRRLTRS
IDIFSMGCVF YYILSKGKHP FGDKYSRESN IIRGIFSLDE MKCLHDRSLI AEATDLISQM
IDHDPLKRPT AMKVLRHPLF WPKSKKLEFL LKVSDRLEIE NRDPPSALLM KFDAGSDFVI
PSGDWTVKFD KTFMDNLERY RKYHSSKLMD LLRALRNKYH HFMDLPEDIA ELMGPVPDGF
YDYFTKRFPN LLIGVYMIVK ENLSDDQILR EFLYS

Possible length 1 = 32 residues longer 673-1115

Possible length 2 = Kinase domain and RNase domain 641-1115

Kinase domain

RNase domain

GST

[P] Prescission Protease cleavage site

The length is either one of the possible lengths mentioned above or just the kinase domain, without any extra length before the GST. See supplementary material of Korennykh et al., 2008.